SICS Technical Report
T98:01

ISRN:SICS-T--98/01-SE
ISSN: 1100-3154

Exploring Navigation; Towards a Framework for Design and Evaluation of Navigation in Electronic Spaces

Nils Dahlbäck, Editor

March 1998

nilsd@sics.se
Swedish Institute of Computer Science
Box 1263, SE- 164 28, Kista, Sweden

Abstract

This document is a slightly revised version of Deliverable 1.1.1 of the PERSONA project, submitted to the European Commission in February 1998. Only typographic errors have been corrected, leading to some changes in the lay-out of the individual pages. The contents and the general lay-out of the individual chapters, including their length are, however, the same as in the original document.

PERSONA is an acronym for PERsonal and SOcial NAvigation. The name of the project illustrates its two-fold approach; studying the individual cognitive, social and cultural differences in navigational ability and recognizing that computer users are social beings in interacting with other people as they make their way through information spaces. Based on this understanding we are developing new approaches to interactive system design. One of these is to identify how and where we can adapt to the individual person's needs. At the same time we are developing alternative approaches to system design, breaking away from the lonely 'walker in the woods' picture of the information system user, to a social being able to interact with other users and so get help in achieving their goals.

In this first deliverable from the project, we present a comprehensive review of literature which we see as having an impact on navigation in information space. This volume contains a number of individual and co-authored papers covering various aspects of geographic and electronic spaces and on navigation in geographic and electronic spaces; Individual and cultural differences; Social aspects of navigation; Design based on alternative or complimentary approaches that we believe hold the promise of making interfaces and systems more navigable.

Keywords : Navigation, social navigation, individual differences, cultural differences, design, electronic spaces, information spaces

Further copies of this document and more information on the PERSONA project can be found at http://sics.se/humle/projects/persona/web/index.html

Exploring Navigation;
Towards a Framework for Design and Evaluation of Navigation in Electronic Spaces

Nils Dahlbäck
Editor

25637 PERSONA
Task 1.1
Deliverable 1.1.1
February 10 1998

Contents

Preface iii

1. Introduction: A Framework for Information Space, Personal and Social Navigation 1
Kristina Höök, David Benyon, Nils Dahlbäck, Rod McCall, Catriona Macaulay, Alan Munro, Per Persson, Marie Sjölinder, Martin Svensson

2. On Spaces and Navigation in and out of the Computer 13
Nils Dahlbäck

3. Beyond Navigation as Metaphor 31
David Benyon

4. Spatial Cognition and Environmental Descriptions 45
Marie Sjölinder

5. Individual Differences in Spatial Cognition and Hypermedia Navigation 59
Marie Sjölinder

6. Social Navigation 73
Martin Svensson

7. Inhabiting Information Space: Work, Artefacts and New Realities 89
Alan J. Munro

8. Route Guidance Issues; Verbal versus Map Instructions, and Route Choices 115
Kristina Höök

9. Evaluating Adaptive Navigation Support 139
Kristina Höök and Martin Svensson

10. Voices in the Forest: Sounds, Soundscapes and Interface Design 159
Catriona Macaulay, David Benyon and Alison Crerar

11. A Comparative Study of Digital and Cinematic Space with Special Focus on Navigational Issues 173
Per Persson

12. Supporting Navigation in Digital Environments: A Narrative Approach 189
Per Persson

13. Navigation In Graphical User Interfaces 199
Rod McCall

Preface

PERSONA is an acronym for PERsonal and SOcial NAvigation. The name of the project illustrates its two-fold approach; studying the individual cognitive, social and cultural differences in navigational ability and recognizing that computer users are social beings in interacting with other people as they make their way through information spaces. Based on this understanding we are developing new approaches to interactive system design. One of these is to identify how and where we can adapt to the individual person's needs. The adaptations can be both dynamic, as in adaptive interfaces, or static, i.e. designing systems or interfaces for clearly defined user groups, or making it possible in for the user to adapt the system to her needs. At the same time we are developing alternative approaches to system design, breaking away from the lonely 'walker in the woods' picture of the information system user, to a social being able to interact with other users and so get help in achieving their goals.

In this first deliverable from the project, we present a comprehensive review of literature which we see as having an impact on navigation in information space. The wide cross-disciplinary team which we have on the project has enabled us to bring together an exceptional body of work. This leads us to a broad framework for understanding navigation in information spaces. This will form the basis of an international workshop on the topic to be held in march 1998. This volume contains a number of individual and co-authored papers covering

- Various aspects of geographic and electronic spaces and on navigation in geographic and electronic spaces: Sjölinder on *Spatial cognition and environmental descriptions*; Dahlbäck *On Spaces and Navigation in and out of the Computer*, and Benyon on *Beyond Navigation as Metaphor.*
- Individual and cultural differences: Sjölinder on *Individual differences in spatial cognition and hypermedia navigation*; and Macaulay, Benyon and Crerar in *Voices in the Forest*
- Social aspects of navigation: Svensson on *Social navigation*, Munro on *Work: Practices, artefacts, design.*
- Social aspects of navigation: Svensson on *Social navigation*, Munro on *Work: Practices, artefacts, design.*
- Design based on alternative or complimentary approaches that we believe hold the promise of making interfaces and systems more navigable: Macaulay et al on sound in interfaces in *Voices in the forest*,, Persson on the use of cinematic methods in *A Comparative Study of Digital and Cinematic Space with Special Focus on Navigational Issues*, and on narratives in *Supporting Navigation in Digital Environments: A Narrative Approach*, Höök and Svensson on adaptive hypermedia in *Evaluating Adaptive Navigation Support,* and Höök on route guidance systems in *Route Guidance Issues*. Evaluation of navigation is addressed in McCall's *Navigation In Graphical User Interfaces*

In the *Introduction* the different perspectives are summarized and presented in an integrating framework for navigation, that will provide the base for further work in the project.

This document is a slightly revised version of the deliverable submitted to the European Commission in February 1998. Only typographic errors have been corrected, leading to some changes in the lay-out of the individual pages. The contents and the general lay-out of the individual chapters, including their length are, however, the same as in the original document.

Chapter 1

Introduction: A Framework for Information Space, Personal and Social Navigation

Kristina Höök, David Benyon, Nils Dahlbäck, Rod McCall, Catriona Macaulay, Alan Munro, Per Persson, Marie Sjölinder, Martin Svensson

We outline a framework for classifying information spaces that we call the "space of spaces". Based on this, we then discuss navigation, personalised as well as socially-enhanced navigational tools. Theoretical issues from a number of different scientific fields are reviewed, and issues on design for navigation are discussed from a number of different perspectives.

This paper also serves as an overview of the other contributions in this volume.

Introduction:
A Framework for Information Space, Personal and Social Navigation

Kristina Höök, David Benyon, Nils Dahlbäck, Rod McCall, Catriona Macaulay, Alan Munro, Per Persson, Marie Sjölinder, Martin Svensson

SICS,
Box 1263
S-164 28 Kista, Sweden

{kia, nilsd, perp, marie, martins}
@sics.se

Dept. Computer Studies
Napier University
219, Colinton Road
Edinburgh, U.K.

{d.benyon, rmccall, c.macaulay,
a.munro}@dcs.napier.ac.uk

http://www.sics.se/humle/projects/persona/web/

Abstract

We outline a framework for classifying information spaces that we call the "space of spaces". Based on this, we then discuss navigation, personalised as well as socially-enhanced navigational tools. Theoretical issues from a number of different scientific fields are reviewed, and issues on design for navigation are discussed from a number of different perspectives.

This paper also serves as an overview of the other contributions in this volume.

INTRODUCTION

Navigation seems to be a fundamental corner stone in many human-computer interfaces. We navigate among our files in our hierarchical file systems, we navigate in large information spaces such as the Internet, we navigate in MUD (Multi-User Dimensions) or VR (Virtual Reality) systems moving from "room" to "room", etc. The purpose of this chapter in the PERSONA literature survey is to provide a framework for discussing various tools for navigation, both personalised and social tools, as well as discussing various aspects of the spaces where this navigation happens. This attempt at a framework will draw upon the definitions and discussions of all the other contributions to the literature survey, and will attempt to place those in a context. We shall start by providing some insights into what we mean by information space, and then discuss navigation in these spaces.

THE SPACE OF INFORMATION SPACES

We take the viewpoint that the concept of "information space" covers a wide range of spaces ranging from the "real" world, via augmented reality[1], VR, to hypermedia, hierarchical file systems, and databases, etc. These spaces vary along several different scales, and these variations will influence how easy or difficult it is for people to access, navigate, find pleasure, or work in them. The dimensions will furthermore influence what sorts of navigational or

[1] Augmented reality is when an information space is superimposed on the real world, as for example when we put up street signs or when we use route guidance systems in the car (see Höök, 1998, this volume). See also discussion in Dahlbäck, 1998, this volume.

orientation tools can be made available to users in the space. In summary, the dimensions we explore are:

- Euclidean properties
- Presentation of objects/relationships/people in the space
- Tasks allowed and performed in the space
- Pre-defined structure of the domain contained in the space
- Size and stability of the space
- Presence of other users and scalability
- Type of organisation of the space (spatial, social, semantic or narrative)
- Moderated versus non-moderated spaces
- Density, distribution, activity, occlusion, and accessibility
- Conceptualisation of objects/relationships/people in the space
- Activity space and information space

Let us discuss each of these dimensions in turn.

Euclidean properties

In the "real" world, certain properties are more or less stable, which means that we can learn about the distances between locations, or relative positions of landmarks, and what we have learnt will not change (at least not rapidly). Stockholm will always be in Sweden – it will not move even if it might grow. This means that users of spaces with this property can make use of distance, direction and relative position. This is true not only for the real world, but works also, more or less in immersive VR, as well as augmented reality. In other spaces, such as hypermedia, relative position is not defined in terms of Euclidean distances, but will instead rely on other ways of measuring distance, relative position, etc. (see discussion in McCall, 1998, this volume). These "made-up" measurements can be similar to distance (number of nodes traversed, etc.) but are different in that they must be learnt specifically for each of these spaces, and their organisation is not fundamental to our cognition in the same way Euclidean distances are. For a more elaborate discussion on Eucledian distances, turn to (Dahlbäck, 1998, this volume).

Presentation

Some spaces have a richer representation that may draw upon visual, auditory, and tactile properties, while others are poorer, and will only rely on abstracted representations that must mostly reside in the user's mind. For example, in UNIX, users do not get any (visual, auditoriy or tactile) feedback on which file catalogue they are in at the moment unless they use a specific command to ask for the information. This means that users have to keep the structure of the file catalogues in their mind and remember where they are right now. Through externalising some of this information onto the interface (as shown in the work by Vicente and Williges, see (Höök, 1998 this volume)), users may be relived some of the burden of keeping track of where they are.

There are several ways by which information can be presented. As argued by Macaulay and colleagues (this volume), sound is one important channel that so far has been little used in the design of spaces and navigational tools.

As pointed out by Sjölinder (1998a, this volume) the space can be presented from three different perspectives: a *survey description* takes a perspective from above, *a route* *description* takes the reader on a mental tour through the environment, and a *gaze tour* locates objects relative to other objects from a fixed outside point of view.

Tasks allowed and performed in the space

In some spaces, there is not much we can do, while others allow for a range of activities. For example, in a database space, we can search for information but not much more, while in a word processor we can produce information, organise it, etc. So in some spaces, navigation will be the most important activity undertaken by users, while in other spaces, navigation is just one aspect of what is going on.

Structure of the domain contained in the space

As discussed in (Dahlbäck, 1998, this volume), we must distinguish between three levels of organisation: the inherent structure of the domain, the structure imposed by the system designers, and finally the cognitive map the user has of the two and their relationships. In some spaces, it will be easy to make the domain structure also organise the whole space as presented to users of it, e.g. in an augmented space, the psychical location of objects can be used as places where the information associated with those locations will be superimposed. In other domains, it is not possible or desirable to make the two structures into one. Some domains might only have a very weak inherent structure, e.g. a collection of food recipes.

In (Persson, 1998a, this volume) the use of narratives as a means to organise and present information is proposed. Such an approach only works for domains that lends themselves to being structured in terms of a narrative.

In general, the domain properties are crucial in understanding how to organise an information space. Those properties are of course linked to why we approach the space, and the same domain may be organised in several different ways depending upon tasks, users, etc.

Size and stability of the space

The sheer size of an information space also determines how much it can be pre-processed or visualised in various ways (e.g. Ahlberg and Shneiderman, 1994). The stability of the space, how often it changes, is another important factor. Given a small, stable space, it is easy to invent maps, personalised guided tours, or have agents present the contents in an interactive way. But if the space is very large and keeps changing and very little can be known of how different parts of the space are and will be related to one-another, tools will have to look quite different.

Concerning Shneiderman's proposal that visualisation solves all our problems with information overflow, we would like to point at the fact that it relies on an assumption that the visual channel always is the most important one. But if we, for example, consider people who live in a jungle, they do not rely on sight – do not think in terms of overviews etc. They 'navigate' by attending visually to what is available visually, aurally to what is available aurally, using memory and narrative to construct spaces outwith the reach of senses or never before encountered. Pre-occupation with visualising means we want to try and take out temporality as an aspect of experiencing space. While when we experience a space through e.g. sound or via a narrative we are also experiencing it temporally. See discussion in Macaulay et al.

Number of users present

In some space, we are on our own. Still other users might have left information behind that we make use of. Munro (1998, this volume) argues that the organisation of a library is something given to us from people who have tried to make our search task easier. Svensson (1998, this volume) argues that indirect social navigation, for example, utilising social filtering, is another way other users may leave traces behind that helps us understand and navigate information spaces.

In some spaces, though, other users are more directly present together with us. Examples are MUD environments such as Palace, OnLive Traveler, etc. so-called inhabited spaces. Munro (this volume) discusses some properties of such spaces and how co-presence of users may enable/enhance subtle communication and coordination between users navigating the space of actions. Munro also discusses the issue of scalability when the space or number of users grows. This issue is also somewhat discussed by Svensson (this volume).

Type of organisation of the space (spatial, social, semantic or narrative)

In Svensson (this volume) and Persson (1998a, this volume) the organisation of a space is discussed. Drawing upon the definition by Dourish and Chalmers (1994), they find that spaces can be organised in three different modes: spatial, social, or semantic (this is also discussed by Dahöbäck (this volume)). Some systems embrace the *spatial* paradigm and structure the information according to some geographical, 'real space' notion. Another 'space' parameter that is attracting interest is *social navigation,* where the movement from one item to another is cued by the activity of another or a group of users. Probably the most common structure, however, is semantic organisation, where the objects in the environment are related through some semantic connection like *similar, alike, more/less general, associated.* To these three Persson (1998a, this volume) adds a fourth, namely a *narrative* organisation mode. The idea is that the nodes/islands of information are connected through a story that is communicated to the user.

This division may be likened to the differences in our cognitive understanding of instructions, as discussed by Sjölinder (1998a, this volume): semantic encoding and spatial encoding.

Moderated versus non-moderated spaces

Navigation – or rather the possibilities for supporting navigation – is strongly influenced by the extent to which the space has a coherent design, or whether it has grown without any control or moderation. In the former case, often called moderated spaces, it is possible for the careful designer to create cues helping the user to orient herself in the space, and being aware of the relationship between the present view and other locations. (See Persson, 1998b, this volume, for some interesting suggestions from cinema on this.) In non-moderated spaces, where no one is on control of the design of the individual nodes or places, nor of their interrelations, this is possible. In this case, navigators must construe their own understanding of the space, and the only possibility for doing so is often retrospective in nature. We should therefore be open to the possibility of design solutions only being applicable to one of these cases.

Density, distribution, activity, occlusion, and accessibility

Chen (1996) provide five different dimensions of spaces that describe how the objects are related to the space:

- density: amount of objects in the world

- distribution: spatial array of objects in the world
- activity: presence and mobility of objects in the world
- occlusion: visibility of objects in the world
- accessibility: whether subjects can reach the destination freely or if they are impeded in some manner.

Conceptualisation versus perception

But the presentation of a space only gives half the answer, as it is not concerned with the interpretation of the space. Missing form the traditional geographies is the failure to appreciate how environments are *conceived* by people as opposed to simply *perceived* by people, as discussed by Benyon (1998, this volume).

Activity space and information space

In Benyon (this volume) a distinction between activity and information space is made. Activity spaces, are the 'real' spaces of physical action, information spaces are those information systems (signs, etc.) that helps us plan, control, monitor and maintain the activity space[2]. We deliberately use the word 'view' instead of 'type' or 'kind' here, to emphasise that activity and information space are intrinsically related. In a socially inhibited world there is never one without the other.

The concept of and information space can now be understood by appealing to the notion of an information system – people, processes, events, activities, and information artefacts. An information artefact is "any artefact whose purpose is to allow information to be stored, retrieved, and possibly transformed" (Green and Benyon, 1996). They all employ symbols to provide information, and in doing so constrains and defines an information space. Thinking about information artefacts means concentrating on particular aspects of the object or device. You have to focus on what information is provided by the artefact and how that information is presented. We can also arrange information artefacts in a hierarchy of different viewports, i.e. windows to a particular sub-space or perspective of the information space.

NAVIGATION

Given our outline with the "space of spaces", we now have a better understanding for all the different situations in which navigation will happen, and the restrictions / possibilities that the properties of the space will give. But what do we mean by navigation? In the work presented here, we investigate two different approaches. One where we more narrowly define navigation (see also the discussion in Dahlbäck, this volume). The second approach takes a more overall perspective on navigation as a means to approach human – system communication in general, as discussed by Benyon (this volume).

Let us start by discussing the first approach.

NAVIGATION AS A SEPARATE ACTIVITY

Navigation is an *activity* undertaken by users in information space. This activity can be broadly divided into four different parts (following Downs and Stea 1973):

---

[2] Note that there is an interesting connection between the activity Vs information space and Dahlbäck's (this volume) action Vs interaction distinction. One is about the 'worlds', the other about the 'agents', but they seem to share at least large parts of an underlying perspective.

- orienting oneself in the environment,
- choosing the correct route,
- monitoring this route, and
- recognising that the destination has been reached.

This division covers a wayfinding activity, but as discussed by Benyon and Höök (1997), navigation is also comprised of *exploration* and *object identification*. When the destination is not known beforehand, the user is exploring the space. In that case, the "destination" in the definition of Downs and Stea above can be thought of as an overall goal to explore the space. The user can undertake this exploration with a more or less definite goal in their minds, be it just for the pleasure of wandering in the space. In those cases, orientation will still be a relevant activity, as will monitoring the path travelled. In these cases there will not be a "correct" route. Recognising that the destination has been reached will in an exploration task more be a matter of feeling that one has had enough. Sometimes exploration will turn into wayfinding, and vice versa.

In an object identification activity, the user is interested in the objects in the space, what they are, their layout relative one-another, or relative the overall space, etc. Again, out of the four activities outlined by Downs and Stea, orientation and monitoring of the route will be most interesting, while choosing the "correct" route or recognising the destination are less important.

As noted by Höök and Svensson (this volume) a lot of research has focused on the activity of helping the user to choose the "correct" route. More efforts should be spent on the layout of the space to allow for exploration and object identification. Orientation, recognition that the destination has been reached, as well as monitoring the route, are equally important activities. Sometimes a richer environment, including sound, tactile feedback, visual feedback, etc., would provide the user with more environmental clues supporting these activities, see (Macaulay et al, Munro, this volume).

Deciding on what the destination (or goal) is can be seen as a process going on between the user and the space (augmented with any tools to support the navigation). When seeing/hearing/feeling parts of the space the user can get a better grip of his/her own desires. As discussed in activity theory (Cole, 1996, Nardi, 1996), the tools mediate the users actions and the two should be studied together. Also, it is in the interaction between users, situations and space, that users attach meaning and understanding of space.

What is particular to navigation, that makes it different from other problem solving activities or in general interaction or dialogue with any system, is the activity of *moving* between locations/nodes (see discussion in Dahlbäck, 1998, Persson, 1998b, Svensson, 1998, all in this volume). Obviously, this will be an important aspect of many different kinds of interactions with the whole spectrum of information spaces (as discussed above). We navigate the action space of an application (see Benyon, this volume), we navigate a hypermedia structure, we navigate in a MUD environment, etc. But we also perform other tasks, such as building objects (documents, artefacts, etc.), solve problems (such as programming), etc., and in those tasks, navigation is but one aspect of our overall activity.

Sometimes it will be important that the user remembers the space after having used it, and then navigation entails recognition.

In summary, we view navigation as an *activity* where we *move* between *locations* / nodes. This activity may be subdivided into four different parts: orientation, route planning, monitoring the route, and recognising the destination. When the destination is known, the navigational activity is wayfinding, while exploration and object identification are navigational activities where the goal is not one particular destination.

So tools that support navigation should consider all these different activities. They should enhance the *quality* of the user's goal, be it getting to a destination, exploration or object identification, as well as *the pleasure / delight* induced by navigating the space (see also Höök and Svensson, this volume).

BEING LOST

So navigation is fundamental to our interaction with many computer applications as well as other information spaces. Unfortunately, navigation might be quite difficult: we get lost, it may induce spatial anxiety (see Sjölinder, 1998b, this volume), we sometimes cannot find whether the destination is even present at all in the space, some spaces keeps changing (such as the web), etc.

Furthermore we know that some people are ”more lost” than others (for a discussion on this, see Sjölinder, 1998b, this volume). These differences stem from individual differences (age, sex, cognitive abilities, personality traits) where some may be culturally dependant (see Sjölinder, 1998b, this volume, and Cole, 1996). Cultural differences in this context can be seen as individual differences arising out of different cultural (and socio-historical) experiences. As expressed by Frake (Frake, 1997):

”Culture provides principles for framing experience as eventful in particular ways, but it does not provide one with a neat set of event-types to map onto the world. [...] Culture does not provide a cognitive map, but rather a set of principles for map-making and navigation. Different cultures are like different schools of navigation designed to cope with different terrains and seas. " (pp 44-45)

From a socio-cultural perspective we might also consider the view that there is no such thing as a stable entity ‘the individual’. Rather we might consider that individuals, cultural groups and differences are defined by context. Using this lens, human behaviour can only be understood in relation to its context (Cole 1996). Part of that context will arise from the socio-cultural-historical setting within which 'we' are situated. For example, feminist geographers have long considered the complex relationships between gender and space/place (Massey 1994).

AN EXTENDED VIEW ON NAVIGATION

This far we have, in line with current work by other researchers in the field, viewed navigation in electronic spaces metaphorically. We have just shown that there are a number of dimensions on which various geographic and electronic spaces differ, and which designers and evaluators of information system need to consider in their work. But instead of analysing the navigation metaphor, we can take a larger leap. If we go beyond the metaphor of navigation (Benyon, this volume), we see a potential for a radical re-thinking about human-computer interaction. Just as the paradigm shift in software engineering from a structured, functional approach to an object-oriented approach has brought significant changes to this discipline, so we may see a change in HCI. We shift from Direct Manipulation to Navigation in Information space. Direct manipulation places the user outside the computer system, distant from the domain. Navigation of information Space places the user inside the system, able to interact more naturally and more directly with the domain rather than the interface. We expect the philosophical implications of such a change to become clearer during the next phase of the project when we build and evaluate concrete examples of personal and social navigation and where we will look at the ramifications for both design and use of interactive systems.

SOCIAL AND PERSONALISED NAVIGATION

If we look at how users navigate in the real world, in cities, in finding information in libraries, in finding their way through buildings, in choosing which television program to see, in deciding which medical doctor to see, etc., we see that this is often done through *talking to other people*. For example, Streeter and Vitello (1985), found that travellers found their way to a destination by driving to the nearest gas station and asking for help. Jon O'Brien (pers.comm.) found that employees in an organisation found it more useful to search for persons than documents, despite having a nicely designed three-dimensional information space of all the documents. Andreas Dieberger (1997) analyses behaviour on the WWW in terms of following other peoples' advice and choices, for example through studying their bookmarks and assembled links. We name this strategy "social navigation", since it does not have to involve building a (mental) spatial model. Instead, social navigation relies on interactions with other people (for a more thorough discussion on this, turn to Dahlbäck this volume). Sometimes it means interacting with another person, sometimes it means studying what a large group of users does (see (Svensson, this volume) on direct and indirect social navigation).

There are several reasons why social navigation may be preferred over more spatially demanding methods. When we talk to someone else, the information we get back can be personalised to our needs. We are perhaps told a little bit more than exactly the information we asked for, or if the information provider knows us, the instructions may be adapted to fit our knowledge or assumed reasons for going to a particular place. The instructions are also adapted to the user in another sense in that they start from the point where the user is at, and is given in a sequential form (first go there, then go there…). Social navigation also has another quality, in that we can judge to what extent the directions given can be trusted depending upon the credibility of the information provider. If I am provided with information on how to find a particularly good definition of some concept from a well-known researcher in my field, I will probably follow the advice, while if the information is given by someone not in my field, I might not even bother to look it up. Yet another aspect of social navigation is the form in which instructions are given, it will frequently be given in a verbal, dialogue-driven form, rather than an abstracted spatial form (for a discussion on verbal versus pictorial/map descriptions turn to Höök, this volume).

Following groups of people, for example, going through an airport may also be considered as a form of social navigation, even if it does not pertain all the properties mentioned above. It does involve the matter of trust. It also provides the seeker with a sense of security: if all these people have chosen this route, it must be the right one. Awareness of other people and their actions may also be communicated and understood in a more subtle way as discussed by Munro (this volume).

The ideas for social navigation conform to what is known about cognition and interaction (see Sjölinder, 1998a, 1998b this volume).

DESIGN BASED ON SOCIAL AND PERSONAL NAVIGATION IDEAS

In the next phase of PERSONA we aim to design and implement some of the ideas around social navigation. Let us briefly present some of these design aspects.

Social navigation will not replace the need for well-designed information spaces and navigational tools that assists users in forming models of the space. It would be stupid to replace alphabetical order in a library with chaos just because there is a librarian that knows the way around the chaos, whom a book borrower can talk to. Furthermore, for some applications, the main task of the user might in fact be to understand the layout and relationships in the space. Our goal is instead to broaden our view on the design space and include social navigation as one tool in the repertoire.

Underlying our design approach is that navigation should be a delightful experience, that part of navigation is goal formulation, and that we need to recognise the risk of making users anxious about getting lost or cognitively overloaded.

A design based on social and personal navigation ideas could take several different paths, e.g.:

- We can try to implement agents that provides users with advice and directions just as other people would (discussed by Svensson, 1998 this volume, and by Persson, 1998a this volume)
- We can design spaces so that we enable communication between users (as discussed by Svensson, this volume)
- We can draw upon some aspects of what makes social navigation successful and use those to build tools (discussed by Munro, this volume)
- We can adapt navigation using intelligent user interface techniques (discussed further in Höök and Svensson, this volume) (see also the discussion on individual differences in (Sjölinder, 1998b this volume)
- Awareness of other peoples activities, as discussed by Munro (this volume)
- Narrative organisation, as discussed by Persson (1998a this volume)
- Off-screen space, as discussed by Persson (1998b this volume)
- The use of soundscape for orientation, as discussed by Macaulay et al. (this volume)

Entertainment and navigation

The focus on narratives and entertainment is based on several tenets as well as reservations (Persson, 1998a, 1998b this volume). Trying to make the navigational experience more emotional, aesthetic or 'story-like' is of course a design end in itself. The, until recently, very 'toolish' conception of computers within the industry and public, is now more an more being replaced by the entertainment industry's notion of computers as 'pleasurable objects' (games, MUDs, chat environments etc.). But perhaps a turn to narratives and interface characters will also support those users with poor spatial ability, since story actions and emotional reactions will possibly be remembered better (than, for instance, spatial relations between landmarks). For sure, narratives and emotional experiences will disturb and hinder a user who has a clear goal and only want to use the computer as a tool. But not all users and situations are like that (think of the *wayfinding*, *exploration*, *identifying objects* distinction made in Benyon and Höök, this volume) We use computers for different purposes. Narratives and entertainment should not be conceived of as an alternative to traditional search engines and navigational tools, but as a parallel and optional alternative for particular occasions.

SUMMARY

We have attempted to place the contributions of the individual writers in this literature review in a context. The structure obtained, is our initial attempt at a framework for navigation.

REFERENCES

Ahlberg, Christopher, and Shneiderman, Ben (1994) Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays, *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems*, 1994.

Benyon, David (1998). Beyond the Navigation Metaphor, this volume.

J. X. Chen, Physically-based Modeling in Information Categorization, *Siggraph'96 Technical Sketches*, New Orleans, Louisiana, August 1996.

Cole, M. (1996). *Cultural Psychology. A Once and Future Discipline*. Cambridge, MA, The Belknap Press of Harvard University Press.

Dahlbäck, Nils (1998). On Spaces and Navigation in and out of the Computer, this volume.

Dieberger, Andreas (1997). Supporting Social Navigation on the World-Wide Web, *International Journal of Human-Computer Studies*, special issue on innovative applications of the Web, Vol. 46, pp. 805-825.

Dourish, P. & Chalmers, M. (1994). Running out of Space: Models of Information Navigation, paper presented at *HCI'94* (British Computer Society).

Downs, R. & Stea, D. (1973). Cognitive Representations, in Downs, R. & Stea, D (eds.), *Image and Environment*, Chicago: Aldine (79-86).

Frake, C. O. (1997) Plying Frames Can Be Dangerous in Cole, M., Engestrom, Y. and Vasquez, O. (eds.) *Mind, Culture and Activity: Seminal Papers from the Laboratory of Comparative Human Cognition*, Cambridge University Press, Cambridge.

Green, T. R. G. and Benyon, D. R. (1996) The skull beneath the skin; Entity-relationship modelling of Information Artefacts. *International Journal of Human-Computer Studies* 44(6) 801-828.

Höök, Kristina (1998) Route Guidance Issues, this volume

Höök, Kristina, and Svensson, Martin (1998) Evaluating Adaptive Navigation Support, this volume.

Catriona Macaulay, David Benyon and Alison Crerar (1998) Voices in the Forest: Sounds, Soundscapes and Interface Design, this volume.

Massey, D. (1994). *Space, Place and Gender*. Cambridge, Polity Press.

McCall, Rod (1998) Navigation In Graphical User Interfaces, this volume

Munro, Alan (1998) Inhabiting Information Space: Work, Artefacts and New Realities, this volume.

Nardi, B. A. (1996) *Context and Consciousness: Activity Theory and Human-Computer Interaction*. Nardi, B.A. (ed.). Cambridge, MA: The MIT Press, 69-102.

Persson, Per (1998a) Supporting Navigation in Digital Environments: A Narrative Approach, this volume.

Persson, Per (1998b) A Comparative Study of Digital and Cinematic Space with Special Focus on Navigational Issues, this volume.

Sjölinder, Marie (1998a) Spatial Cognition and Environmental Descriptions, this volume.

Sjölinder, Marie (1998b) Individual Differences in Spatial Cognition and Hypermedia Navigation, this volume.

Streeter, Lynn A. , Diane Vitello, and Susan A. Wonsiewicz (1985) How to Tell People Where to Go: Comparing Navigational Aids, *Int. J. Man-Machine Studies*, 22: 549-562.

Svensson, Martin (1998) Social Navigation, this volume.

Chapter 2

On Spaces and Navigation In and Out of the Computer

Nils Dahlbäck

SICS

Current work on navigation in electronic worlds is based on the assumption that geographic and electronic worlds are similar enough to make it possible to use results from work on environmental psychology and related areas in the design of electronic information spaces. The present paper is an attempt to analyze the underlying assumptions behind this approach in some detail, as well as an attempt to describe a number of different dimensions on which these spaces can differ. We also discuss how these differences might influence user behavior and design.

On Spaces and Navigation In and Out of the Computer

Nils Dahlbäck
SICS, Box 163, S164 28 Kista, Sweden[1]
nilsd@sics.se
http://www.sics.se/humle

Current work on navigation in electronic worlds is based on the assumption that geographic and electronic worlds are similar enough to make it possible to use results from work on environmental psychology and related areas in the design of electronic information spaces. The present paper is an attempt to analyze the underlying assumptions behind this approach in some detail, as well as an attempt to describe a number of different dimensions on which these spaces can differ. We also discuss how these differences might influence user behavior and design.

Introduction

We are 'lost in hyperspace' when 'navigating electronic worlds'. And many workers in HCI are working on remedies for this. Many issues of relevance for this task were addressed at a workshop on navigation in electronic worlds at the CHI'97 conference in Atlanta. As pointed out by Jul and Furnas (1997) in their summary of the workshop, "no definitive solutions were reached, [but] much of the problem space was laid out." Our aim with this paper is to contribute further to the task of clarifying the problem space.

As pointed out by Lakoff and Johnson (1980), each metaphor hides more than it highlights. The risk of being blinded by an inappropriate use of a metaphor makes it is therefore important to also assess the pros and cons of the metaphors used. In the present case, it seems especially important to clarify what is distinct and unique to navigation, in contradistinction to other tasks performed by users, as well as the similarities and differences between geographic space and different kinds of electronic spaces.

The Navigation Metaphor

The navigation in electronic worlds-metaphor rests on some important assumptions. First and foremost that geographic worlds and electronic worlds are similar enough to make it possible to make use of results obtained by researchers on navigation and wayfinding in geographic space. "One approach to the problem, which we have found to be beneficial, is to compare navigation in the physical world with navigation in electronic worlds" (Hirtle, 1997, p. 1). Another, and less obvious underlying assumption, is that the *activities* of wayfinding and navigation are similar to the information seeking activities of users of information spaces.

A corollary of these basic assumptions is that it is assumed that navigation in geographic space is similar enough to navigation in electronic worlds to make it possible to use results obtained by geographers, spatial psychologists, architects and others in designing computer systems that can be navigated with ease. But all electronic worlds are not created equal. And there are possibilities for more than one kind of activity in these spaces. Since it is not obvious that navigation support will look the same in all these cases, it seems important to attempt to

[1] Also with the Department of Computer and Information Science, Linköping University

clarify which the important dimensions distinguishing the different categories are. Without doing so, it will be more or less impossible to be able to generalize from particular design solutions or results from empirical studies to the relevant group or class of similar instances.

Definitions of navigation and wayfinding

The Marriam-Webster dictionary defines 'navigation' and 'navigate' as follows (abridged).

1. NAVIGATION

1: the act or practice of navigating

2: the science of getting ships, aircraft, or spacecraft from place to place; especially: the method of determining position, course, and distance traveled

3: ship traffic or commerce

- nav·i·ga·tion·al /-shn&l, -sh&-n&l/ adjective

- nav·i·ga·tion·al·ly adverb

2. NAVIGATE

Etymology: Latin navigatus, past participle of navigare, from navis ship

+ -igare (from agere to drive) -- more at AGENT

Intransitive senses

1: to travel by water: SAIL

2: to steer a course through a medium; specifically: to operate an airplane

3: GET AROUND, MOVE

Transitive senses

1 a: to sail over, on, or through b: to make one's way over or through: TRAVERSE

2 a: to steer or manage (a boat) in sailing b: to operate or control the course of (as an airplane)

In recent work on navigation in electronic spaces, navigation has been used as a synonym for 'wayfinding', a term used in architecture. Wayfinding has been defined by Passini (1984, p 154) as "a person's ability, both cognitive and behavioral, to reach spatial destinations", which is only one, but perhaps the most central, of the aspects of navigation described in the dictionary. This activity is composed of a number of analytically distinct but in actual acting intertwined sub-parts. Downs and Stea (1973) claim wayfinding to be composed of four steps: *orienting* oneself in the environment, *choosing* the correct route, *monitoring* this route, and *recognizing* that the destination has been reached.

At a recent workshop on Navigation in Electronic Worlds (Jul and Furnas, 1997) a number of leading workers in the field presented their definitions of navigation. In some cases these definitions are clearly about, or at least heavily influenced by, activities in geographic space, e.g. Leventhal "Navigation is the cognitive process of acquiring knowledge about a space, strategies for moving through space, and changing one's metaknowledge about a space". Others, e.g. Spence, focus on activities in hypermedia spaces only; "I have suggested that navigation broadly comprises four activities: browsing, context modeling, gradient perception and movement." A similar view is put forth by Dourish and Chalmers (1994), who define navigation as "the means by which a user can describe movement between pieces of information".

For some workers in the field, 'navigation' seems to have acquired a meaning synonymous with 'seeking information', as illustrated in the following quotation (Wexelblat,1997) "Imagine that you borrow a book or paper reprint from a colleague. It comes to you not as it would from a store, but rather with pages folded, with notes in the margins, possibly with tabs or tags attached to mark interesting pages. These additional features allow you to *navigate* the book in a different way that if it were untouched" (italics added). Such a wide definition of 'navigation', making it more or less synonymous with information retrieval, is not what is aimed for here. To be able to e.g. provide guidelines on navigational aspects of design we need to make clear what sets navigation apart from other tasks of users. If not, these guidelines will be co-extensive with general guidelines for good HCI design. On the other hand, the state of the art in the field today is not developed enough to make it meaningful or even possible, to define the concept on navigation in electronic worlds or information spaces in a universally accepted way. Further theoretical and empirical work is needed to reach this goal

We need both to consider to which extent geographic space is similar to different kinds of electronic spaces, and to consider to which extent the activities pursued in electronic spaces are similar to navigation. We will begin with the issue of different kinds of spaces, and return later to the issue of the different kinds of activities.

Geographic and electronic spaces

While most workers in the field stress the similarities between geographic and electronic worlds, it is as important to also stress the differences. One important such difference is that geographic space has a stable Euclidean geometry, making spatial relations between objects stable and permanent. Gothenburg will always be between Stockholm and Edinburgh. To a large extent this is true for VR systems, and especially immersive VR systems. This is not, however, true in a hypertext or hypermedia system, where new links can arbitrarily be created, making previously distant nodes adjacent to each other. This difference is emphasized also by Kent Wittenburg, who claims "The concept of navigation in cyberspace has a completely different physics from navigation in the physical world. The cartographers and engineers of cyberspace must not only create the maps and the instruments but also the world itself. In such a plastic medium, it is not clear that navigation and search should be thought of as contrasting approaches, but rather as part and parcel of the same activity" (1997, p 1).

But, as pointed out above, cyberspace is not one but many. Today's common classification distinguish between hypertext and hypermedia, and immersive and non immersive Virtual Reality (VR). It seems clear that different design solutions are required for all these kinds of electronic worlds. 3D Virtual Reality (VR) systems, and especially so-called immersive ones, preserve most properties of the geographic worlds, hereby making the mapping from work in the geographic world to the electronic world rather straightforward. But, as pointed out by Wittenberg above, for hypertext and hypermedia, the mapping is less perfect, and more caution should be exercised when trying to make use of results obtained by workers on navigation in the physical world, as well as when trying to port successful design solutions from one kind of electronic space to another.

The standard classification is, however, primarily technical, and does not consider the *contents* of the information space, nor the *tasks* performed in it. Leaving the task aspect aside for the moment, it seems clear that there are different levels of structure in most systems of this kind, and this is especially true for hypertext and hypermedia systems. We can at least distinguish between three levels that need to be kept distinct; the *inherent structure* of the content domain, the *structure imposed* by the system designers, and finally the *cognitive map* that the user has of these two structures and their interrelationships. When it comes to the structure of the system, it is probably necessary to distinguish between the underlying structure and the presented structure. In an ideal world, and at least for simple enough systems, these will be co-extensive.

When developing designs or evaluating or comparing existing designs, the kind of mapping between the structure of the information presented and the structure of the information presentation need to be considered. At least the following cases of hypertext information content-structure mappings can be distinguished.

- Information about geographic and similar information, e.g. tourist information about the hotels at a summer resort. Here there exists a real spatial structure that can be used by the users for structuring the information, and furthermore some of the information is inherently spatial in nature, e.g. the distance from the hotel to the beach.
- Information about domains which are not real physical spaces, but which have some commonly agreed upon internal structure. Examples of this would be biological classification systems or educational systems (which often are described in terms of 'higher' and 'lower' education etc.).
- Information about domains that do not have any commonly agreed upon internal structure. Examples of this are classification of different kinds of musical styles or classification of art.

It should be noted that for the last category there might exist consistent conceptual structures for sub-groups of users. And many competing such structures or worldviews can co-exist and even compete at the same time. In fact, the difference between the second and third category is very relative to a particular cultural perspective. All people that have received traditional Western schooling would probably agree upon at least the crude outlines of a biological classificatory system, but the Kaluli people on Papua New Guinea would probably not share this view. So in some sense we have not three but two distinct types. But the distinction between the latter two has an heuristic value when designing hypermedia systems for specific purposes, in making the designer forced to consider whether there exist a structure common to all the intended users, or if many different such structures need to be catered for in the design.

For the first and second category, the users will probably have some shared cognitive map of the inherent structure, which system designers should stay close to when designing the interface and the underlying structure of the system. The third category requires extra caution, since there will be no mutually shared cognitive map of the inherent structure, probably requiring a very clear mapping between the system's underlying structure and the structure presented by the interface.

Another interesting classification is the one by Dourish and Chalmers' (1994) of three major modes of navigation, namely *spatial*, *social*, and *semantic*. Leaving the social navigation aside here, spatial organization, the prime example being immersive VR systems, is obviously closely related to our ability to navigate in geographic space. The class called 'semantic' (CD-ROMs, help systems, etc), which has a structure organized by semantic connections, is argued by Persson (this volume) to rely on the user's semantic and not spatial ability. The distinction is obviously important, but some data (also reviewed by Persson) suggest that further clarification of what characterizes the different classes might be needed here. It has been shown that hypermedia, databases, and hierarchical file systems are of a spatial character (Dahlbäck et al, 1996, Benyon & Murray, 1993, Vicente & Villeges, 1988). But these systems are not spatial in the same sense as an immersive VR system. They seem rather to belong to the non-spatial or non-geographical system in the tripartite classification presented above. They have, in a sense, a spatial structure but not spatial content. The concept of time seems to be a prime example of this (Clark, 1973, see also Dahlbäck 1992). There are a number of different conceptual structures of time, in different cultures or for different purposes (linear, circular, and subtypes of these), but they all seem to have in common the use of a spatial structure to conceptualize something inherently non-spatial.

In this context, it is interesting to note the results from a study by Stanney and Salvendy (1995). It has been previously shown that users with high spatial ability outperform users with low spatial ability (Dahlbäck et al, 1996, Benyon & Murray, 1993, Vicente & Villeges, 1988). Stanney and Salvendy showed that it was possible to build interfaces that compensated for the differences between high and low spatial users, but that the differences between the two groups remained when the task required the construction of a mental model of the information.

Some observations by Dahlbäck et al (1996) in fact suggest that different human spatial abilities are correlated with navigation in geographic space and navigation in a large help system, but more work is clearly needed here. It is important in further work to clarify whether these two kinds of spatiality share enough properties to make similar solutions work in both cases, or whether they should be treated differently, despite both being of spatial nature. I would like to suggest here that a distinction need to be made here between *relational* knowledge and *object* knowledge need to be upheld here. This is similar to Piaget's (e.g. Piaget & Inhelder, 1973) distinction between *operational* (i.e. structural), and *figurative* knowledge, and to Shum's (1990) distinction between *locational* and *attributional* information (though Shum claim both of these to be sub-categories of spatial information, making the much too common mistake of conflating 'visual' and 'spatial').

Electronic spaces usually lack one important feature of geographic space, namely the explicit or implicit information that we are progressing in the right direction. This is not only given us by route signs telling us that we have less distance left to our final destination. When walking in the forest in search for a good place to stay the night, preferably by some lake or river, we are presented with an abundance of cues possible to use for monitoring our task; the slope of the hill and whether we are walking up or down, but also the changes in the kinds of flowers growing on the ground, the kind of soil we are walking on etc., help the experienced hiker to find a way towards a suitable place for staying over night. Both Furnas (1997) in his work on *navigational residue* and Pirolli in his work on *scent* (1997) have addressed this problem (though the work of Pirolli and his co-workers is concerned with the task of summarizing and communicating the structure of very large collections of information). Both approaches have shown initial promising results, but much further work is clearly needed here. Another way of managing the lack of environmental awareness, is to incorporate features in the interface that makes it easier for the user to connect the different snap-shots provided into a coherent real or virtual space (c.f. the chapter by Persson (this volume) on cinematic space).

Unfortunately, many design solutions for alleviating the navigation difficulties encountered by users are limited to *moderated* spaces only (i.e. databases, help systems, and similar designed systems). For *non-moderated* spaces (the prime example being WWW), it is not possible to through design create a uniform perspective or a tailored navigational instrument like a map. In these cases metadata needs to be derived, making the navigational support task a much larger challenge.

Navigation and other activities

The most obvious conclusion possible to draw from staying close to he navigation metaphor is that the major stumbling block for the user is finding the site or location of the relevant information. In the vast majority of cases, our main concern is finding the place we are looking for. Once we have found it, the navigation task is solved. But perhaps this is taking the navigation metaphor too literal? Sometimes this is not the task the user is engaged in. Höök and Benyon (1997) point out that not all activities of computer users are navigation or wayfinding in the strict sense of the word. They make a distinction between *wayfinding* and *exploration*. The former refers to the situation where the navigator has a clear and precise goal or task, the latter refers to situations where such a clearly defined goal does not exist, or where the user only wants to obtain a general overview of the spaces. There is a difference, then, between *learning* the space and *using* the space. In actual practice, these two activities are of course always intertwined. The reason for making a distinction between them in the present case is that it seems plausible that the obstacles encountered by the user when engaged in two activities are different, and that therefore different kinds of user support might be needed.

Furthermore, in those cases where the user is not engaged in exploration of the space, but in using it to solve a particular task, the major problem seems not always to be finding the specific locations in the information space. In their work on the design of a database for research geneticists, Doerry and co-workers have found that, while users have little trouble finding specific data, they frequently become distorted during multi-step data manipulations (Doerry et al 1977). So, at least in this case, the problem is not finding particular pieces of information, but rather the structuring of the information found. Staying with the navigation terminology, the problem is not finding the route to the goal, but construing the survey map of the domain.

Doerry et al also points out that the structure of the interaction with the information system is rarely related to the user's actual task. (C.f. below in the section on spaces). Another interesting observation by Doerry et al is that users often switch between different tasks, making it very difficult to infer the appropriate task structure from the interactions with the system

Kinds of navigation in information spaces

As pointed out above, Benyon and Höök (1997) have argued for making a distinction between wayfinding and exploration. Furnas (in Furnas & Jul 1997) takes this one step further. He distinguishes between two tasks, *searching* and *browsing*, and two tactics, *querying* and *navigation*. These are defined as follows (abridged here). Searching; looking for a known target, Browsing; looking to see what is available, Querying; submitting a description of the object sought to a search engine, Navigation; moving around sequentially in an environment, basing the next step on the task and current information about the environment.

Navigation and problem solving

In some cases, the models of navigation used in the HCI community, (e.g. the one developed by Spence (1997) (see fig 1)) show large similarities to general models of problem solving developed within the AI community. This seems somewhat problematic. Not only is this model in many, perhaps most, cases not a correct characterization of problem solving activities as pointed out by Suchman 1987). What is more important in the present context, is to make clear what the characteristics unique to navigation are, that makes this task separate from problem solving in general. The *situatedness* seems to be one such important factor (c.f. Furnas above). Navigation in the full sense of the word is never an armchair activity; it always involves the locomotion through space. Because of this, there is always the possibility of revising the activities planned during the task (the *monitoring* aspect of the Downs and Stea classification). An important consequence of this, is that error recovery and other forms of adaptation become an intrinsic part of the activity. For us as designers it becomes important to look at navigation design with this in mind. Perhaps support for error recovery, and more general, for monitoring the progress during locomotion, is more important than providing the electronic world analogues of maps and compasses? It is clear that present day interface design gives little support for this.

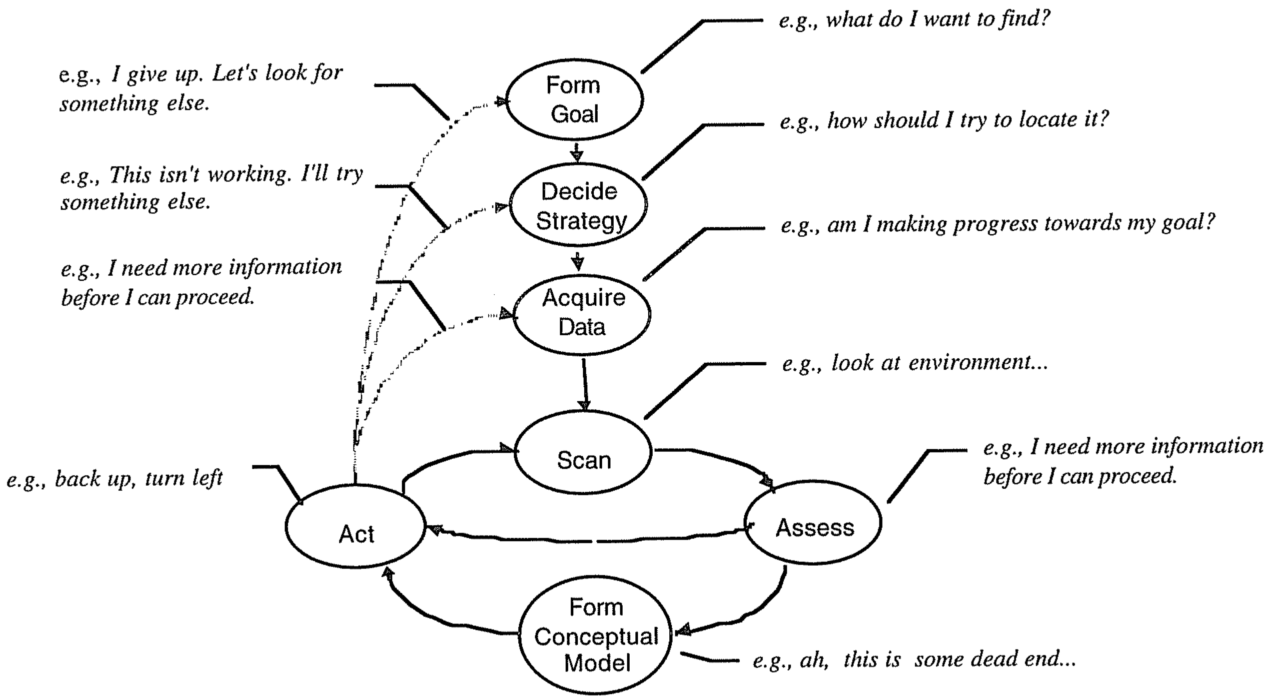


Figure 1 A General Framework for the Navigation Process. From Jul and Furnas (1977)

Another aspect important to navigation addressed both by Persson and Svensson (this volume) is that navigation is a cognitively active process. They distinguish navigation from *transportation*, where the subject either knows the path extremely well, or when the subject is just following a path that he knows will take him to the destination, but whiteout knowing, or caring to know, how the start and end points are situated in relation to each other, or which places he will pass during the locomotion through space. From a psychological point of view, the two kinds of transportation are probably different. In many cases of e.g. routine commuting by car the driver often have a good survey knowledge of the space, and can switch to another mode should the need arise, e.g. in the case of a traffic accident blocking the normal route. This is true for many other highly trained or automated tasks, which makes them very different from the passive transportation of the passenger in the bus, who presumably cannot take over the drivers task should the need arise.

At the CHI'97 workshop on navigation a similar characterization was made. In their report from the workshop, Furnas & Jul (1997) describes it as follows:

> After some discussion of the individual definitions, there were four basic aspects upon which everyone could readily agree. While they do not constitute a definition, they will likely be essential to one:
>
> *Locomotion*
> Something moves—either the navigator or an object that is the focus of the navigator's attention. This assumes a concept of *location*, in particular, a *here* and a *there* (or not-here). The movement is *directed*, i.e., deliberate decisions are made in choosing among locations. The movement is *purposeful* in that it is undertaken in service of meaningful goal.
>
> *Decision-making*
> In being a directed and purposeful activity, decisions must be made continually regarding strategies for reaching the goal and determining whether the goal has been reached. These decisions sometimes follow a plan and sometimes respond to the environment. They depend on both declarative and procedural knowledge and frequently require coordination of knowledge in different forms (orientation).
>
> *Process*
> Navigation is an incremental real-time process that integrates these two components (locomotion and decision-making).
>
> *Context*
> Each navigation process takes place in a particular information environment (set of locations) and is inextricably tied to that environment.

For those scholars that wish to see navigation as something more specific than problem solving in general, there seem to be a consensus on a view with the following properties: The navigator is an active agent. (A passenger in a bus with darkened windows is no more navigating than the body in a hearse.) There is a movement through a space, and this movement is monitored and adapted during the process. There is a goal for the process, and once the goal is reached the process is finished. If this, then, is what constitutes navigation, the question that emerges is to which extent this is a reasonable description of the activities in electronic worlds?

Learning to navigate in a space[2]

It is common to distinguish between three kinds of knowledge of a space. Following Siegel and White (1975) these are called *landmark*, *route*, and *survey* knowledge, often acquired in that order. Landmarks are conceptually and perceptually distinct locations. Route knowledge is understanding of the environment in terms of paths between locations. Survey knowledge describes the relationships between locations, often likened to map like memory representations. Results from Tversky (1991) and others indicate these survey representations often show a hierarchical structure.

[2] This field is reviewed in depth by Sjölinder (this volume). The text here owes much to her work and my discussions with her on this and related topics.

The fact that humans often naturally create hierarchical survey knowledge structures, is of course very relevant for the designers of hypertext and hypermedia systems, which by definition have this structure. It is, however, important to note some complications here. First, that there probably has to be a close mapping between the structure of the system and of the user's knowledge structure. Second, that the mental map of a domain will vary depending on the viewer's perspective. The latter studies can, however, be interpreted as indicating that it is the metric and not the topological aspects of the mental maps that vary view viewer perspective. Since hypertext and hypermedia system do not have a stable Euclidean geometry, it is possible to entertain the hypothesis that designers need not take these fluid perspective changes into account when designing their systems. But this is, of course, only a hypothesis, in need of further investigation.

Acquisition of spatial knowledge can be either primary or secondary, where primary denotes knowledge acquired through navigation in the world and other kinds of primary experience, and secondary denotes knowledge acquired through symbolically transmitted information, e.g. maps (Schachter and Nadel, 1991).

There are three different ways of describing an environment to a listener; taking the listener on an imagined *tour* through the environment, providing a *survey* description, or taking the listener on a so-called *gaze tour* (Tversky, 1991). The difference between the two latter is that in the first case a map like overview is given, whereas in the latter, the listener is placed on an imagined viewing point, and the locations of the different objects in the world is given with reference to the viewer's position (e.g. the Hoover Tower is to the left of Jordan Hall, the Gates building to the right, and the Tresidder behind it).

It has been shown by Tversky and others, that the knowledge acquired through studying a map is in some senses different from the knowledge of the same space acquired from actually being there. The best-documented difference is that primary spatial memory is more robust to various manipulations of orientation at the time of retention. On the other hand, a series of studies by Taylor and Tversky (1992a, b) have shown that subjects acquire the same spatial mental models from survey and route descriptions, as well as from maps.

The interesting question that then emerges from these findings is what kind of knowledge that is acquired through navigating in a hypermedia space. This activity is in some sense similar to being in the space. But what distinguishes it from the actual locomotion through geographic space, is that no cues, or at the best very abstract and map like cues, are given showing how one particular location is related to the others in the space. So in this respect it is more like map reading.

Linde and Labov (1975) and Taylor and Tversky (1996) have shown that there are typical preferences for different kinds of relations for different kinds of environments. When describing an apartment, most people prefer a route description, whereas for describing rooms that can be seen from one viewpoint, most people prefer gaze descriptions. But there are also individual differences here; e.g. some people use survey descriptions to describe their apartments. To this author, this suggests that navigational aids should be able to adapt to the user's individual preferences in environmental description format.

Navigation tools supporting the user

Much of present day research is devoted to designing tools that will help the user navigate to space, by e.g. giving her a map of the domain. But once again, we can ask ourselves whether we are staying too close to the navigation metaphor, and a particular interpretation of that metaphor, namely the professional navigator in a ship or a plane. Is it really tools we need? Designing maps and similar devices to help the user see where he is in the space, could be likened to having a person drive a window less car through the help of a sophisticated map and navigation system, where a simpler and probably more efficient solution should be to give her a window to look out through?

But looking out through a window is of no help, if everything you see look the same no matter where you are. The suggestions put forth by Persson (this volume) and MacAulaylay, Benyon, and Crerar (this volume) all suggest ways of giving locational cues to the user, without interfering with her primary task. Note, however, that these suggestions in many cases will be in conflict with commonly accepted design guidelines, which advocate a uniform design of each window or section of a large system. This is most likely a contradiction that can be resolved through a careful design, but more work is clearly needed before these design solutions can be formulated as general design guidelines.

In summary; there are two ways of supporting navigation, at least in moderated spaces. One can make the space more navigable, by incorporating in the design features that makes orientation easier, or one can develop tools to support the navigation task, by e.g. presenting map-like overviews of the space.

Social navigation (no user is an island)

Taking one step back, and looking at present day research on navigation, it becomes obvious that the implicit picture of the navigator or information system user, is in most cases that of some kind of lonely 'walker in the woods'. It takes only a moment of reflection to realize that this is a very distorted picture of human activities. We are in constant interaction with other persons through most or all steps in the tasks and activities pursued.

As a reaction to this, a number of different scholars have recently introduced the concept of 'social navigation' (e.g. Dourish and Chalmers 1994, Erickson 1996, Dieberger 1997). As pointed out by Svensson (this volume) the definitions differ between different scholars. The common denominator seems to be an attempt to introduce social aspects into the task of navigating in information spaces. No attempt will be made here to recapture Svensson's review and analysis; the interested reader is referred to Svensson's chapter here. Instead we will discuss the relationship between different interpretations of the navigation metaphor and different views of social navigation. We will also discuss how to delimit the 'social' in social navigation from other kinds of navigation.

Social navigation in the different metaphor interpretations

Analytically, we can distinguish between two ways that social aspects can be incorporated in the activities of users of information spaces. First, the social activities can be likened to getting help from other people in finding the way to the desired goal, e.g. asking people in the street for the way to the library. In this case we preserve the goal of the activities in the metaphor's mapping between the domains, making the important goal the finding of a particular place in hyperspace. Second, the social activities can be likened to asking for information from other people, instead of finding out from written and other sources, e.g. asking people for information, instead of going to the library to find the information there. In the latter case, we are in a sense substituting social interaction for navigation activities.

For the first interpretation, supporting social navigation in electronic worlds implies developing agents and other software devices that the user can use for finding the way to the desired goal. For the second interpretation, supporting social navigation would more likely imply creating software information providers that the user can access through some kind of interactive dialogue, without having to find the way to the information sources, and perhaps in many cases not even having to learn the location of these sources. Another approach would be to support the social interaction and dissemination of information on the Web and in other similar electronic worlds (c.f. Dieberger 1997). Svensson places himself squarely in the first camp "social navigation is navigation. In fact (...) the only thing that separates navigation and social navigation are the tools."

One could, however, argue that this is not an either or situation; that we have to decide which of these two interpretations that is the correct one. We have previously pointed out that there are kinds of activities that users of information spaces pursue that are different from the prototypical navigation task. Also these tasks are always done within a social context, and the social aspects need consequently to be considered in successful design supporting these tasks too.

One should also be aware that the distinction made above is not a dichotomy. We are never individuals interacting with either the social or the physical world. We are all the time doing both, also when we are alone. A good example of this is given by Wexelblat (1997), who points out that following the paths in a forest is actually using information given by other persons; they have walked here because it lead to some interesting or in other respects valuable destination. (The Footprints system currently under development at the MIT Media Lab tries to provide this kind of information for users of the WWW.) Svensson addresses similar issues under the heading of 'indirect social navigation'. One problem here is that is seems difficult to delimit indirect social navigation from navigation in general, since all the electronic and other information spaces that exist are created by people, they are in a sense vehicles of socially transmitted information. This is especially true if we also include so-called history enriched environments. The World Wide Web is, as pointed out by Svensson, essentially a social structure. But it is a strange social structure; created by people, but (with a few exceptions, like the Footprints system) inhibited only by individual users.

Social and non-social navigation

Is there then no way of delimiting social navigation from other forms of navigation? Following Dahlbäck (1998) we could argue that there are two fundamentally different ways of being in the world, *action* and *interaction*. Action is what we are doing when manipulating the physical world; walking, moving or changing objects etc, interaction is always with humans and other agents, i.e. dialogue[3]. Interaction is always with another agent, to whom we can ascribe inten-

[3] It is of course unfortunate that the computer industry has used the words 'dialogue' and 'interaction' for aspects of human's computer usage that in fact are more like analogues of physical manipulation, e.g. graphical user

tional states (c.f. Dennett, 1987). Interaction is always social, action is never social per se, but can of course take place in a social context, and the environment in which most people spend most of their time is to a large extent created by human activities.

We can then use this distinction to make a difference between social and non-social navigation. Social navigation is in essence navigation which at least to some part is done through interaction with other agents, human or artificial. It is what Svensson call 'socially enhanced navigation'.

The basic argument for distinguishing between these two classes is that we as humans carry with us different expectations concerning action and interaction; most of us would be greatly surprised if our car would engage in a dialogue trying to convince us that we shouldn't turn right at the next intersection.

The hypothesis advanced here is, that since we can both act and interact with computers, it becomes important to make clear to the user which state the system is at a particular time. Furthermore, that designs that place themselves on some middle ground between action and interaction should be avoided, especially for computer novices.

Summary: Dimensions we might need to consider

We have surveyed and discussed a number of different dimensions pertaining to spaces and activities in spaces. Let us try to summarize here what we have found. The basic distinction was between geographic and electronic spaces. Among the latter, we discussed a number of different issues.

Electronic spaces

- Hypertext, Hypermedia, Virtual Reality, Immersive Virtual Reality
- Moderated Vs non-moderated spaces

Levels of structure in hypertext and hypermedia systems

- Inherent structure, imposed structure, presented structure, cognitive structure

Kinds of structural/spatial information

- Spatial (geographic), structured but non-spatial; non-structured
- Spatial Vs visual information

Tasks in electronic spaces

- Searching and browsing; querying and navigation
- Finding a piece of information Vs (finding and) structuring many pieces of navigation
- Using a space Vs learning a space

---

interfaces. But that does not invalidate the distinction, it only makes it more difficult to communicate it to that community. Moving a chair by hand, or asking a person to move *are* after all very different kinds of actions.

Kinds of learning of a space

- Primary (experiential) Vs secondary (symbolic)

Social dimensions

- Non-populated Vs populated spaces (social traces in the world)
- Social Vs non-social navigation
- Acting Vs interacting

While not all of these dimensions will be important in all cases, we would like to argue that in both design and evaluation we need to make clear where in this multi-dimensional space we currently are working. If not we risk importing successful design features from other systems that are less useful in our present case, or we risk making overgeneralizations of the results obtained from evaluation studies. No claim is being made here that the dimensions we have considered here constitute an exhaustive list. But we hope that we have provided a useful starting point for further work in this area.

Acknowledgements

Valuable input on the contents of this paper have been provided by the other project members, and is greatly acknowledged, but the final responsibility of the text is by its author, and should not be taken as positions agreed upon by the entire project. Much of the text is clearly influenced by my participation in the CHI'97 workshop on navigation in electronic spaces, and the knowledge and inspiration received from the other participants in this event is greatly acknowledged here. Additional financial support was provided by HSFR (The Swedish Council for Research in the Humanities and Social Science).

References

Clark (1973) Space, time semantics and the child. In T. Moore (Ed.) *Cognitive Development and the Acquisition of Language*. New York: Academic Press.

Benyon and Höök (1997) Navigation in Information Spaces: supporting the individual. In *Proceedings from Interact'97*. Australia, July 1997.

Benyon and Murray (1993) Adaptive Systems; From Intelligent Tutoring to Autonomous Agents. *Knowledge Based Systems*. 6(3).

Dahlbäck, Nils (1992) Representations of Discourse; Cognitive and Computational Aspects. Ph.D. Thesis. Linköping University.

Dahlbäck, Nils (1998) The Navigation Metaphor, How far will it Take us? Manuscript in preparation.

Dahlbäck, Nils, Höök, Kristina, and Sjölinder, Marie (1996) Spatial Cognition in the Mind and in the World; The Case of Hypermedia Navigation. In *Proceedings from the Eighteenth Annual Meeting of the Cognitive Science Society*. San Diego, Ca, July 1996.

Dennet, Daniel C. (1987) *The Intentional Stance*. A Bradford Book, The MIT Press.

Diebergber, Andreas (1997) Supporting Social Navigation on the World Wide Web. *International Journal of Human-Computer Studies*. 46, 805-825.

Doerry, E., Douglas, S., Kirkpatrick, T., and Westerfield, M. (1997) Task-centered Navigation in a Web-accessible Dataspace. In *Proceedings from the AACE/WebNet97 Conference* Toronto, Canada, November 1, 1997
http:/zfish.uoregon.edu/ZFIN/PAPERS

Dourish, Paul and Chalmers, Matthew (1994) Running Out of Space: Models of Information Support. In *Proceedings from HCI'94*. Glasgow, August 1994.

Downs, R and Stea, D (1973) Cognitive Representations. In R.Downs and D. Stea (Eds.) *Image and the Environment*. Chicago: Aldine, pp 79-86.

Erickson, Thomas (1996) The World Wide Web as Social Hypertext. *Communications of the ACM*, January, 1996.

Furnas, George (1997). Effective View Navigation. In *Proceedings from CHI'97*. Atlanta, Ga, 22-27 March 1997, 367-374.

Hirtle, Stephen (1997). Spatial knowledge and navigation in real and virtual environments. Position paper for the CHI'97 workshop on Navigation in Electronic Worlds.
http://www.pitt.edu.~hirtle/navig.html

Jul, S., Furnas, G. (1997). Navigation in Electronic Worlds. SIGCHIBulletin, 29, 4 (Oct). New York: ACM Press.

Lakoff, George and Johnson, Mark (1980), *Metaphors we live by*. Chicago: The University of Chicago Press

Linde, C and Labov, W. (1975). Spatial Networks as a Site for the Study of Language and Thought. *Language*, 51: 924-939.

MacAulay, Benyon and Crerar (1998). Voices in the Forest: Sounds, Soundscapes and Interface Design. *This Volume*.

Passini (1984). *Wayfinding in Architecture*.

Persson (1998a). A Comparative Study of Digital and Cinematic Space with a Special Focus on Navigational Issues. *This Volume*.

Persson (1998b). Supporting Navigation in Digital Environments: A Narrative Approach. *This Volume*.

Piaget, Jean, and Inhelder, Bärbel (1973). *Memory and Intelligence*. London: Routledge and Kegan Paul.

Pirolli, Peter. (1997) Computational Models of Information Scent-Following in a Very Large Browsable Text Collection. In *Proceedings from CHI'97*. 3- 10.

Schachter, D.L. and Nadel, L. (1991). Varieties of Spatial Memory: A Problem for Cognitive Neuroscience. In R.G. Lister andH.J. Weingarter (Eds.) *Perspectives on Cognitive Neuroscience*. Oxford: Oxford University Press.

Shum, S. (1990). Real and Virtual Spaces: Mapping from Spatial Cognition to Hypertext. *Hypermedia*, 2(2): 133-158.

Spence, (1997) cited in Jul and Furnas (1997)

Stanney, K.M. and Salvendy, G. (1995). Information Visualization: Assisting Low Spatial Indivituals with Information Access Tasks through the Use of Visual Mediators.*Ergonomics*. 38(6): 1184-1198.

Suchman, Lucy (1987) Plans and Situated Action: The Problem of Human-Machine Interaction. Cambridge, MA: Cambridge University Press..

Svensson (1998) Social Navigation. *This Volume.*

Taylor, Holly and Tversky, Barbara (1992a). Spatial Mental Models Derived from Survey and Route Descriptions. *Journal of Memory and Language*. 31: 261-292.

Taylor, Holly and Tversky, Barbara (1992b). Descriptions and Depictions of Environments. *Memory and Cognition*. 20(5): 483-496.

Taylor, Holly and Tversky, Barbara (1996). Perspectives in Spatial Descriptions. *Journal of Memory and Language*. 35: 371-391.

Tversky, Barbara. (1981) Distortions in memory for maps. *Cognitive Psychology*, **13**, 407-433.

Tversky, Barbara (1991) Spatial Mental Models. *The Psychology of Learning and Motivation*. 27.

Wexelblat, Alan (1997). Navigation by Trail-following. Position paper for the CHI'97 workshop on Navigation in Electronic Worlds. http://wex.www.media.mit.edu/people/wex/CHI97-wkshp-pos.html

Vicente and Villeges (1988). Visual Momentum as a Means of Accomodating Individual Differences Among Users of a Hierarchical File System. In J. Rasmussen and P. Sunde (Eds.) *Proc 5th Symposium EFISS,* Risö National Laboratory, Denmark, November 1987. New York: Plenum Publishing.

Wittenburg, Kent (1997). Navigation and Search – What's the difference? Position paper for the CHI'97 workshop on Navigation in Electronic Worlds. http://www-personal.engin.umich.edu/~sjul/nav97/wittenburg.html.

Chapter 3

Beyond Navigation as Metaphor

David Benyon

Napier University

The popular view of navigation is that it is a conscious, goal directed activity in which someone is trying to reach a destination. Such a view of navigation is essentially individualistic, objectivist and cognitive. Just as Lynch's aim in trying to specify the basic, discrete units of (urban) space have had a significant impact on urban planning and design, so the metaphorical use of the concept of navigation in information space may lead to effective and important design decisions. However, whether this view of navigation can act as the basis of an understanding of what people do in such spaces is another matter. The crucial thing missing form the traditional geographies and traditional views of navigation is the failure to appreciate how environments are conceived by people as opposed to simply perceived by people. An environment is not simply some physical structure to which humans must adapt. People play a role in producing the space, through their activities and practice. A semiotic analysis of space recognises that there are many different views of space and that space is a subjectively defined concept. There is a context to space which needs to be communicated, negotiated and understood between people. More than just space, there is the idea of place. People produce or construct their places at different times and there is a knock on effect from one place to another. In this chapter some implications of taking this different view of information space are explored.

Beyond Navigation as Metaphor

David Benyon

Dept. Computer Studies, Napier University
219, Colinton Road
Edinburgh, EH14 1JD
d.benyon@dcs.napier.ac.uk

The popular view of navigation is that it is a conscious, goal directed activity in which someone is trying to reach a destination. Such a view of navigation is essentially individualistic, objectivist and cognitive. Just as Lynch's aim in trying to specify the basic, discrete units of (urban) space have had a significant impact on urban planning and design, so the metaphorical use of the concept of navigation in information space may lead to effective and important design decisions. However, whether this view of navigation can act as the basis of an understanding of what people do in such spaces is another matter. The crucial thing missing form the traditional geographies and traditional views of navigation is the failure to appreciate how environments are *conceived* by people as opposed to simply *perceived* by people. An environment is not simply some physical structure to which humans must adapt. People play a role in producing the space, through their activities and practice. A semiotic analysis of space recognises that there are many different views of space and that space is a subjectively defined concept. There is a context to space which needs to be communicated, negotiated and understood between people. More than just space, there is the idea of place. People produce or construct their places at different times and there is a knock on effect from one place to another. In this chapter some implications of taking this different view of information space are explored.

Introduction

If we take the concept of navigation metaphorically, then we begin with notions of sea faring and the navigation of ships and boats. It would be fair to say that navigation is also now extended to navigation on land and in space and more generally to be the activity of finding ones way throughout an environment. In architecture, the term 'wayfinding' is preferred (and used synonomously with the term 'navigation'). Passini (1984, p154) defines wayfinding as, "a person's ability, both cognitive and behavioral, to reach spatial destinations." He bases his conception on Downs and Stea (1973) who see wayfinding as composed of four steps: orienting oneself in the environment, choosing the correct route, monitoring this route, and recognizing that the destination has been reached.

Such a view of navigation is essentially individualistic, objectivist and cognitive and there has been much work on how people develop 'cognitive maps' of their environment which enables them to find their way to a specific location. Early studies of environments such as those conducted by Lynch, culminating in the classic 'The Image of the City' (Lynch, 1960) identified features of the environments; edges, paths, nodes, districts and landmarks. Lynch's aim in trying to specify the basic, discrete units of (urban) space have had a significant impact on urban planning and design, but whether they can act as the basis of an understanding of what people do in such spaces is another matter.

In direct opposition to 'The Image of the City', 'The City and the Sign' (Gottdiener and Lagopoulos, 1986) presents a number of views from urban semioticians that highlight the limitations of the Lynchian and cognitive perspective. The crucial thing missing form the traditional geographies is the failure to appreciate how environments are *conceived* by people as opposed to simply *perceived* by people. Ledrut (1986) indicates the importance of this by arguing that if we accept the objective identification of landmarks, districts and so forth, this does not distinguish between humans finding their way and animals moving through a maze. The environment is not simply some physical structure to which humans must adapt. People play a role in producing the space, through their activities and practice. This view is strongly re-inforced by Lefebvre (1991) in his 'The Production of Space' and others in a post-modern tradition; geographers such as Soja (1993) and de Certeau (e.g. 'Walking in the City', 1993), feminist writers such as Massey (1996) and semioticians such as Barthes (writing on urban semiotics, 1986), Eco (writing on architectural semiotics, 1986) and so on. The cognitive approach to geography leaves the use that people make of their environment out of the analysis. This is not to say that all studies of cognitive mapping and all the analysis provided by cognitive geographies needs to be thrown away. Only that the social construction and the ideological impact of space needs to be considered also. So, Gottdiener (1986) can comment that 'In the case of [shopping] malls...on the one hand the mall is the materialization of the retailers intention to sell consumer goods...on the other hand, the mall is the physical space within which individuals come to a participate in a certain type of urban ambiance'. And Barthes (1986) 'that two neighbourhoods are adjoining, if we rely on the map...while, from the moment when they receive two different significations, they are radically separated in the image of the city'. It is these differences which are ignored by the objectivist tradition of the analysis of space.

The semiotic analysis of space recognises that there are many different views of space and that space is a subjectively defined concept. We recognize that we as people negotiate a shared understanding of space. There is a context to space which needs to be communicated, negotiated and understood between people. More than just space, there is the idea of place. People produce or construct their places at different times and there is a knock on effect from one place to another.

For example, a group of academics had a meeting in the meeting room. This space became a place for them to have a meeting, but only because no-one else had booked it. If they had had the Christmas party in the meeting room it would have become a different place for that period, and if the party were too noisy, the offices nearby would no longer be work places for their occupants. The city centre retains it's physical structure but changes from a shopping place during the day to a place where gangs hang out at night. The gangs have notions of 'their' space which would not be recognised by the shoppers.

With these ideas navigation (as opposed to wayfinding) includes a number of activities which occur in space and which people are willing to agree are sorts of navigation. Thinking about space also leads us to think that virtual spaces, or information spaces, may be like physical ones or they may not be.

We can, then distinguish navigation from wayfinding. Clearly people navigate *through* places, so navigation is not always directed towards a destination. I might be going camping in the highlands and so I navigate through Glasgow. Someone might be trying to get away from something - e.g. a football crowd. In this case navigation is away from a place, looking for somewhere safe. Away from the geographical space and towards a social space - i.e. a space for safety.

Different people see things differently. The shoppers and the gangs see street corners in different ways (Street Corner Society by William Whyte). There are different conceptions of landmarks, districts, etc. depending on race, gender and social group. The ship's captain can see many different landmarks in the ebb and flow of a river than the novice. A climber might tell me about a landmark but I may totally fail to see it.

There is little objectivity in the physical world. Can you really measure the length of a coastline? But you can certainly navigate it. Is the drunk navigating his way home really engaged in a conscious activity? If I drive past my turning on the way home, or make some other 'expert slip' am I conscious? It seems that navigation can be unconscious.

In considering navigation as a metaphor, then, we must be careful not to load ourselves with too much of the metaphor's baggage. Similarly we need to be careful about how previous analyses of the phenomena may have missed important aspects. To quote from Barthes again writing on understanding the urban; 'The city is a discourse and this discourse is truly a language: the city speaks to its inhabitants, we speak our city....The real scientific leap will be realized when we speak of a language of the city without metaphor.'

In a similar vein we may argue that 'Navigation of Information Space' is not (just) a metaphor, it is a paradigm shift. Navigation of Information Space is a new paradigm for thinking about HCI, just as Direct Manipulation was a new paradigm in the 1980s. Shifting the paradigm changes the way you think about things. Computers still compute even though we have an object-oriented paradigm instead of one based on functional decomposition. It happens that people believe that thinking about computing as OO helps us to develop better systems. In this approach, then, we seek the 'real scientific leap' by emptying Navigation of Information Space of its metaphorical meaning so that the real meaning can emerge. If you were to ask people what their image is of people using computers today, they would describe a person looking at a VDU. This image is important because it means that we see people outside the computer, looking in onto the world of information, trying to access systems, trying to find things out. But if you ask those people about the image of other activities such as driving, shopping, or having a meeting, they will describe people as inside the activity, in a social environment, part of and surrounded by the various objects the are using. It is this change from seeing the person outside the information space to seeing the person inside the information space that is the important change of view.

Activity and Information Space

By adopting the concept of an information space, we immediately set up a correspondence with something that is not an information space. We call this the activity space - the space of 'real world' activities (or at least the space of physical action). In order to undertake activities in the activity space, people need access to information. Thus the information space is intrinsically linked with the activity space.

As humans we live, work and relax in information spaces. At one level of description all our multifarious interactions with the experienced world are effected through the discovery, exchange, organisation and manipulation of information. Information spaces are not the province of computers. They are central to our everyday experience. Finding your way through an airport, a hotel or a city involves travelling though the activity space, using an information space. Paper documents represent another type of information space. Users will get quite different information from books, from newspapers and from magazines. Similarly they will find various information in timetables, guides and maps. Interacting with other people involves sharing some activities in an information space.

Our purpose — as information system designers — is to design and develop structures and procedures which will make information available to people which will help people to undertake their activities; to help them plan, control and monitor their undertakings. People undertake some activities using some artefact(s) (thus activities are always mediated by artefacts). Activities are processes which make use of artefacts within the system and are triggered by some events. This system (or activity space) has a boundary, outside of which the activities can be considered to have little or no influence. Events may arise from inside the system or from outside (Figure 1). A system is a coherent set of interdependent components which exists for some purpose, has some stability and can usefully be viewed as a whole[1]. As illustrated below, such systems consist of people, artefacts, processes and the activities which are undertaken as a result of events.

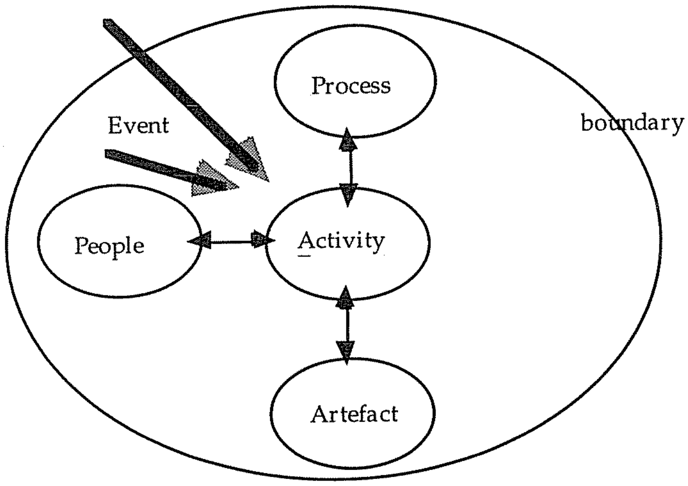


Figure 1 The components of a system; an activity space

Each of these systems also has an information system, the purpose of which is to help plan, control, monitor and maintain the activity space. The information system may be informal (for example, casual conversation), or it may be formalised to a greater or lesser extent.

The information system (or information space) is a system which has a similar structure to the activity system. The important difference is that it uses information artefacts to represent relevant features of the activity system. Users of the information system engage in various activities by performing various processes on the information artefacts — such as selecting items of interest, calculating totals, looking at pictures, etc.. The information space uses signs,

[1] This definition comes from Checkland's work (Checkland, 1981)

structured into information artefacts to represent (certain aspects) of the activity space. This is illustrated in Figure 2.

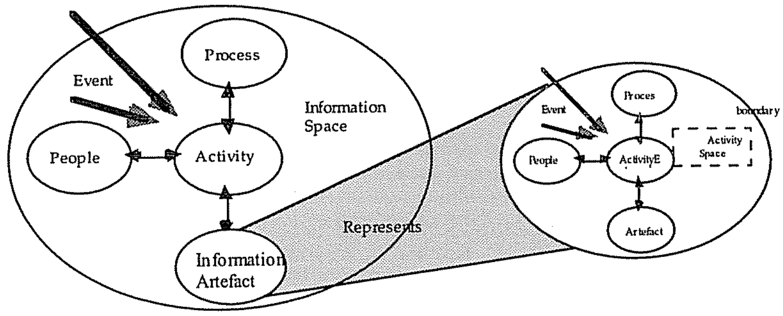


Figure 2 The Information Space

Information Artefacts

The concept of an information space can now be understood by appealing to the notion of an information system - people, processes, events, activities and information artefacts. An information artefact is 'any artefact whose purpose is to allow information to be stored, retrieved, and possibly transformed' (Green and Benyon, 1996). Interactive devices such as spreadsheets, word processors, and music notations are clearly examples of information artefacts, but so are non-interactive devices like tables, documents and musical scores. An information artefact consists of two levels of description; a conceptual level provides some abstraction of the experienced world and a perceptual level provides a view, or viewport onto that structure. Thus a paper train timetable is one information artefact providing information about train journeys and a talking timetable is another (Benyon, Green and Bental, in press).

All information artefacts employ various symbols, structured in some fashion, and provide functions to manipulate the symbols (whether conceptually or physically). From these symbols people are able to derive information. Thus every information artefact constrains and defines an information space - the symbols, structure and functions which enable people to store, retrieve and transform information.

We have posited the existence of two spaces; the activity space and the information space. In the activity space people undertake actions — they drink coffee, travel to Edinburgh, buy clothes and so on. To undertake these activities they use information and carry out activities in the information space. For example, to drink some coffee, I might see a sign, KENCO, which exists in the activity space. I need to locate a button (so undertake and information space activity), press the button (in the activity space), take the coffee (in the activity space) and so on. To do these things I have perceived and interpreted various information artefacts that rely on culturally determined signalling conventions (K-E-N-C-O spells KENCO. KENCO is a sign for coffee because KENCO make coffee).

Information artefacts have a conceptual structure and one or more viewports onto that structure. In the case of simple signs such as KENCO the conceptual and perceptual sides of the information artefact are tightly coupled. The letters denote the word, but the word connotes coffee. In more complex information artefacts the conceptual and perceptual structures may have a more complex relationship. I can view my files in various different ways - by date, by name and so on. Thus there are different viewports onto the same conceptual structure.

Information artefacts may be arranged in a hierarchy of different viewports, each of which gives more detailed information. The Macintosh 'Finder' or Microsoft 'Windows' are information artefacts that reveal the files that are stored on the computer and the functions that the computer can perform. Each function that 'Finder' reveals will typically have another information artefact which reveals information about how to use that function. For example, the existence of a calculator function is revealed by a menu item or icon labelled 'Calculator' (Figure 3). The calculator display in turn reveals more information about the calculator's functions by using a layout which looks very like the layout of a 'traditional' hand-held calculator, and which includes icons for specific functions such as '+', '-', '*' etc. These specific functions are not explained any further by information artefacts; instead the user is expected to know that (for example) '*' means 'multiply', and the user is expected to understand the effects of the multiply function without any further help. The calculator icon information artefact allows the user to double click in order to open the calculator. The icons on the calculator itself behave very differently if double clicked.

When we consider an information artefact, we need to distinguish between two levels of description. We need to consider both the information that the artefact presents to us and the way in which the information is presented. More formally, we need a conceptual description (or conceptual model) of the information in the artefact and a description of the perceptual display (or perceptual model) of that artefact. For example, we could have a conceptual model of a clock as a device which represents time in hours and minutes. This conceptual model could have a variety of perceptual displays, such as an analog clock face, a digital clock face, or a speaking clock. Different viewports enable different tasks to be performed more or less easily.

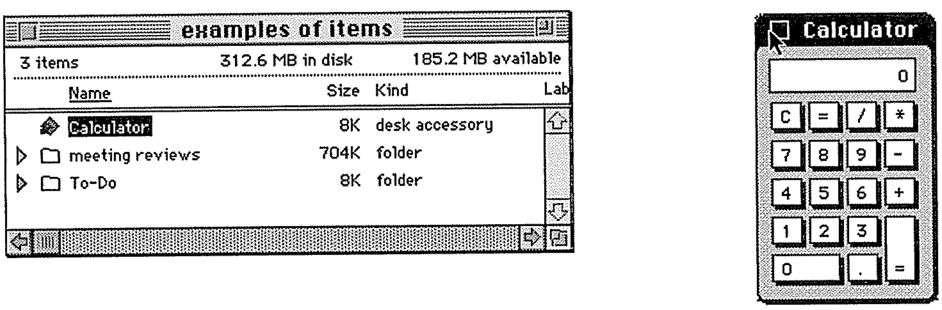


Figure 3 Hierarchical viewports onto information structures: the Macintosh Finder reveals the calculator icon; launching the calculator reveals another interface.

An information artefact abstracts certain aspects of an object or device and then employs some perceptual device to reveal that abstraction to the user. In a well designed information artefact, the abstraction which is chosen highlights the important features of the object or device. Ideally, the abstraction should also be closely related to concepts that the user already knows about, and the perceptual device used to reveal the conceptualisation should capitalise on the user's existing knowledge. For example, the Macintosh 'Calculator' offers similar operations, and uses a similar layout, to that of a familiar pocket calculator.

An information artefact consists of:

- a conceptualisation of objects in the experienced world which has the purpose of revealing some information about the underlying objects to some users (and possibly allowing changes to be made).
- a viewport which provides access to that conceptualisation and which employs a method of presentation from which the user may derive information (and if necessary provides means for making changes).

Just as there may be several levels of viewport onto the same underlying structure, so there may be different viewports at the same level. The same information can be presented in different ways; some methods of presentation may hide certain information altogether, or they may make it quicker and easier to discover some kinds of information than others. The telephone network can be viewed in terms of subscribers who have a name, address, exchange and telephone number. The viewport onto this is usually provided by a telephone book which lists the relevant information. The telephone book presents the data in alphabetical order of subscriber name which affects the usability of this view - so searching for a particular number, not knowing the subscriber's name, becomes a nearly impossible task. Other viewports onto this structure (e.g. the 'Directory Enquiries' or the 'Call Tracing' viewport) enable other goals to be achieved.

A train timetable could be presented on paper or as a talking timetable. The different methods of presentation may look very different, but since the underlying structure of the information is the same many aspects of the different information displays will also be the same.

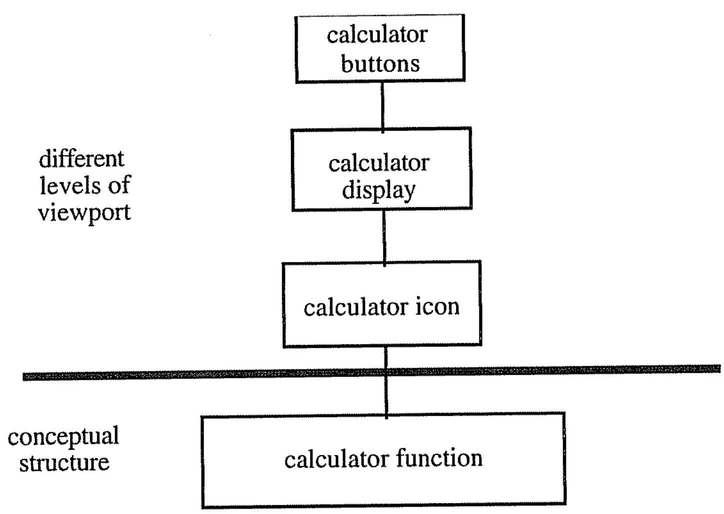


Figure 4 Illustrating a hierarchy of viewports for a calculator information artefact

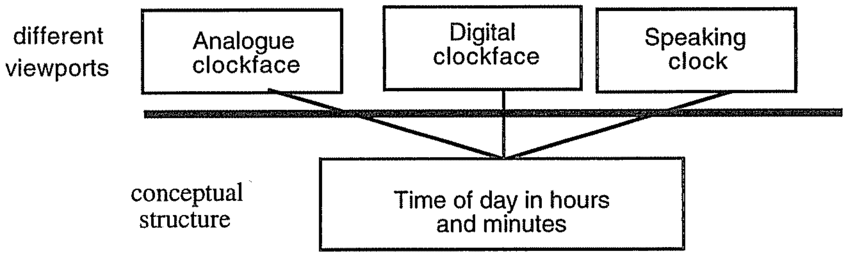


Figure 5 Illustrating different viewports at the same level of abstraction

A talking timetable typically gives only a summary of the departure time from one major station and the arrival time at the main destination (i.e. it hides some information), whereas a paper timetable usually gives times at the intermediate stops. With a talking timetable it can take substantially longer to find out when the evening trains leave, since the times are usually read in sequence – i.e. it is harder to discover some kinds of information.

We have seen that an information artefact consists of;

- a conceptualisation of some object(s) in the experienced world which has the purpose of revealing some information about the underlying object(s) for some users
- a viewport which provides access to that conceptualisation and displays the actual data from which the user may derive information.

However, whenever we create a perceptual display (a viewport) it then becomes an object in the experienced world (the activity space). Consequently it may have its own information artefact designed to reveal information about the display. This leads to the hierarchical arrangement described above.

The fact that there can be multiple levels of information artefacts, each built upon the others is important. The process may be seen as follows. Some designers recognise something about the experienced world. They choose to conceptualise the structure of this in a particular way and develop a conceptual model. In order to access this conceptualisation, various viewports are created. Each viewport reveals something about the underlying structure and it is this combination of viewport and conceptual structure which we call an information artefact. The viewports can now be considered in a similar fashion –we recognise that the viewports are part of the experienced world. Accordingly we can conceptualise the viewport and provide a view onto that conceptual structure. This produces another information artefact.

In theory this can continue indefinitely with information artefacts being created on top of one another revealing different aspects of the underlying experienced world. In practice, users tend to specialise in a particular level of discourse about information artefacts and lower levels effectively disappear from their experienced world. For example, one view of the world of computers deals with the physical arrangement of files on discs, with the workings of disc access times and transfer rates, memory allocation and so on. For most of us such a view is not experienced, instead it is presented to us through information artefacts such as computer operating systems. The graphical user interfaces which are so ubiquitous provide another view onto the operating system and so the experienced world for us is one of dragging icons, double clicking and using menus to issue commands. The car mechanic experiences a different world from the driver. The surgeon experiences a different world from the patient. Recognising these different levels of information artefact is important in order to establish a shared level of abstraction within which we can discuss our needs and concerns.

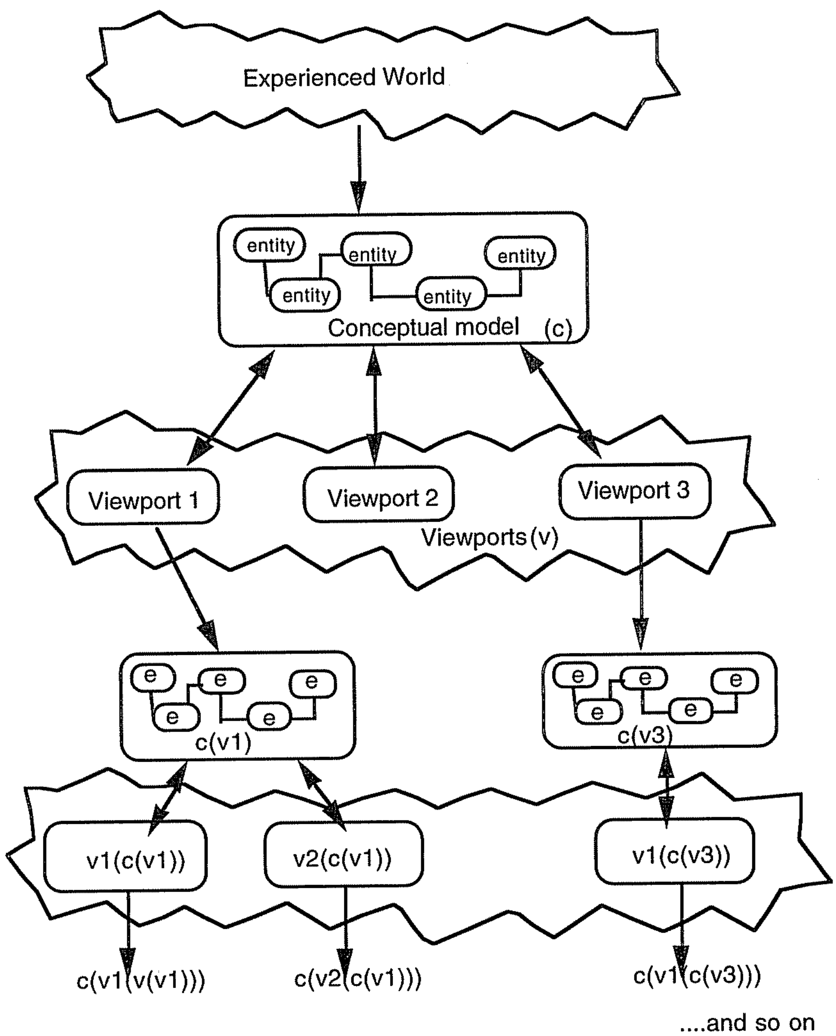


Figure 6 When we create viewports they become part of the activity space

Conclusion

With this conception, navigation in information space is exactly the activities which people undertake in the information space. Navigation in information space means manipulating the information artefacts, moving between levels of viewport and between different viewports. When using a word processor, for example, the actual typing of the document's contents is not an information space activity, but as soon as the user looks at the text and sets about to reposition the cursor, change the font, or select a menu item, so they step immediately into the information space. In the information space they must perceive and interpret the information artefacts so that they can achieve their activity space goals.

As information space designers, we need to consider how people will interpret the information artefacts; how they will conceive the underlying domain from their perceptions of the form and function of the viewports. Eco (1986) comments that in architecture, architectural objects are not there primarily to communicate, they are there to function. Thus a roof is clearly intended to keep out the rain and a doorway functions to let people go into or out of a building. In a similar way my word processor provides various functions to format my text. But by recognising that these are information artefacts, we recognise that their form (the viewport that we provide onto the conceptual structure) denotes their function. Recognising something as a roof allows the person to get shelter from the rain, recognising something as a doorway allows us to get into the building and recognising 'Font...' allows us to format our text.

In producing information spaces, designers create information artefacts to denote functions. But once people enter the information space they interpret not just the primary function of the information artefact. They will make other connotations about the culture, history and ideology that surrounds, and is defined by, that space. The full meaning of, and activities undertaken with, information artefacts are not determined by the designer: they are produced by the users. We need, then, to populate our information spaces with information artefacts that enable and encourage people to understand the activity space. We do that by facilitating navigation in the information space.

References

Barthes, R. (1986) Semiology and the Urban in Gottdiener, M. and Lagopoulos, A. Ph. (Eds.) (1986) *The City and the Sign*. Columbia University Press, New York

Benyon, D. R., Green, T. R. G. and Bental, D. (1998)*Conceptual Modelling for Human Computer Interaction, using ERMIA* Springer, London

de Certeau, M. (1993) Walking in the City In During, S. (Ed.) *The Cultural Studies Reader*, Routledge, London

Downs, R. & Stea, D. (1973). "Cognitive Representations", in Downs, R. & Stea, D (eds.), *Image and Environment*, Chicago: Aldine (79-86).

Eco, U. (1986) Function and Sign: Semiotics of Architecture in Gottdiener, M. and Lagopoulos, A. Ph. (Eds.) (1986) *The City and the Sign*. Columbia University Press, New York

Gottdiener, M. and Lagopoulos, A. Ph. (Eds.) (1986) *The City and the Sign*. Columbia University Press, New York

Gottdiener, M. Recapturing the Center: A semiotic analysis of shopping malls In Gottdiener, M. and Lagopoulos, A. Ph. (Eds.) (1986) *The City and the Sign*. Columbia University Press, New York

Green, T. R. G. and Benyon, D. R. (1996) The skull beneath the skin; Entity-relationship modelling of Information Artefacts. *International Journal of Human-Computer Studies* 44(6) 801-828

Ledrut, R. (1986) Chapter 5 In Gottdiener, M. and Lagopoulos, A. Ph. (Eds.) (1986) *The City and the Sign*. Columbia University Press, New York

Lefebvre, H. (1991) *The Production of Space* Blackwell

Lynch, K. (1960) *The Image of the City*. MIT press, Cambridge, MA

Massey, D. (1996) *Space, Place and Gender* Blackwell, Oxford

Passini, *Wayfinding in Architecture*

Soja, E. (1993) History: geography: modernity. In During, S. (Ed.) *The Cultural Studies Reader*, Routledge, London

Chapter 4

Spatial cognition and environmental descriptions

Marie Sjölinder
SICS

The concept of spatial cognition or spatial ability captures several quite different cognitive functions. One way of cluster spatial tests, according to certain aspects of the spatial ability, will be presented. Difficulties in measuring or capturing different aspects of the spatial ability will also be discussed. Further this review describes issues concerning the acquisition and representation of spatial knowledge. Finally there will a presentation of different environmental perspectives and spatial descriptions.

Spatial cognition and environmental descriptions

Marie Sjölinder
SICS Box 1263 S-164 28 Kista Sweden[1]
marie@sics.se
http://www.sics.se/humle/

The concept of spatial cognition or spatial ability captures several quite different cognitive functions. One way of cluster spatial tests, according to certain aspects of the spatial ability, will be presented. Difficulties in measuring or capturing different aspects of the spatial ability will also be discussed. Further this review describes issues concerning the acquisition and representation of spatial knowledge. Finally there will a presentation of different environmental perspectives and spatial descriptions.

Introduction

The word spatial is defined as "concerning or existing in space" (Oxford Advanced Learner's Dictionary of Current English, 1992) or "relating to space" (Collins Dictionary of the English Language, 1986). This review will discuss spatial cognition - how human beings deal with issues concerning relations in space, navigation and wayfinding. A brief evolutionary perspective on spatial cognition will be presented. Different spatial tests will be discussed and clustered together with respect to different aspects of the spatial ability and different perspectives on the acquisition of spatial knowledge will be presented. The description of an environment is dependent on the viewer's perspective and of the characteristics of the space to be described. Some different environmental descriptions will be described, and differences between these viewpoints will be discussed.

Spatial cognition

Spatial abilities are cognitive functions that enable people to deal effectively with spatial relations, visual spatial tasks and orientation of objects in space. One aspect of these cognitive skills is spatial orientation, which is the ability to orient oneself in space relative to objects and events; and the awareness of self-location (Reber, 1985). It might be hypothesised that spatial abilities were even more crucial for survival in prehistoric than in contemporary times. At least survival in current urban living conditions requires different cognitive abilities which involves more verbally based abilities. Hunting people with specific ecological demands usually present good visual discrimination and excellent spatial skills, for instance the embedded figures test is better performed by cultural groups for whom hunting is important for survival. Nonetheless, contemporary city humans spatial abilities are not necessarily inferior to prehistoric humans and may have evolved with the new living and cultural conditions (Ardila, 1993). The right hemisphere specialisation for spatial abilities can be correlated with language acquisition and evolution. Spoken language evolved with the appearance of new cultural conditions, and one hypothesis is that spatial abilities also evolved with the appearance of new cultural and environmental conditions. It might be proposed that prehistoric humans presented a more bilateral representation of spatial abilities. Not only language development and complexity, but also the development of new spatially based abilities may have increased the right hemisphere specialisation for handling information with spatial content, as well as the left specialisation for linguistic abilities (Ardila, 1993).

---

[1] Also with the Department of Psychology, Stockholm University, and the Department of Computer and Information Science, Linköping University

The spatial abilities are a part of the fluid intelligence, when dividing a general intelligence into subgroups by making a distinction between fluid intelligence and crystallised intelligence. Fluid intelligence is usually assessed with tests of general reasoning ability, but also with math and spatial ability measures. Measures of fluid intelligence reach a peak in early adulthood and then decline slowly, accelerating in decline after the age of 60. Crystallised intelligence represents the sum of acquired knowledge and experience and it is usually measured with tests of verbal knowledge and general information. Measures of crystallised intelligence typically show an increase until late middle age, then remain stable (with the possibility of continuing increases) through the age of 80 (Rolfhus & Ackerman, 1996).

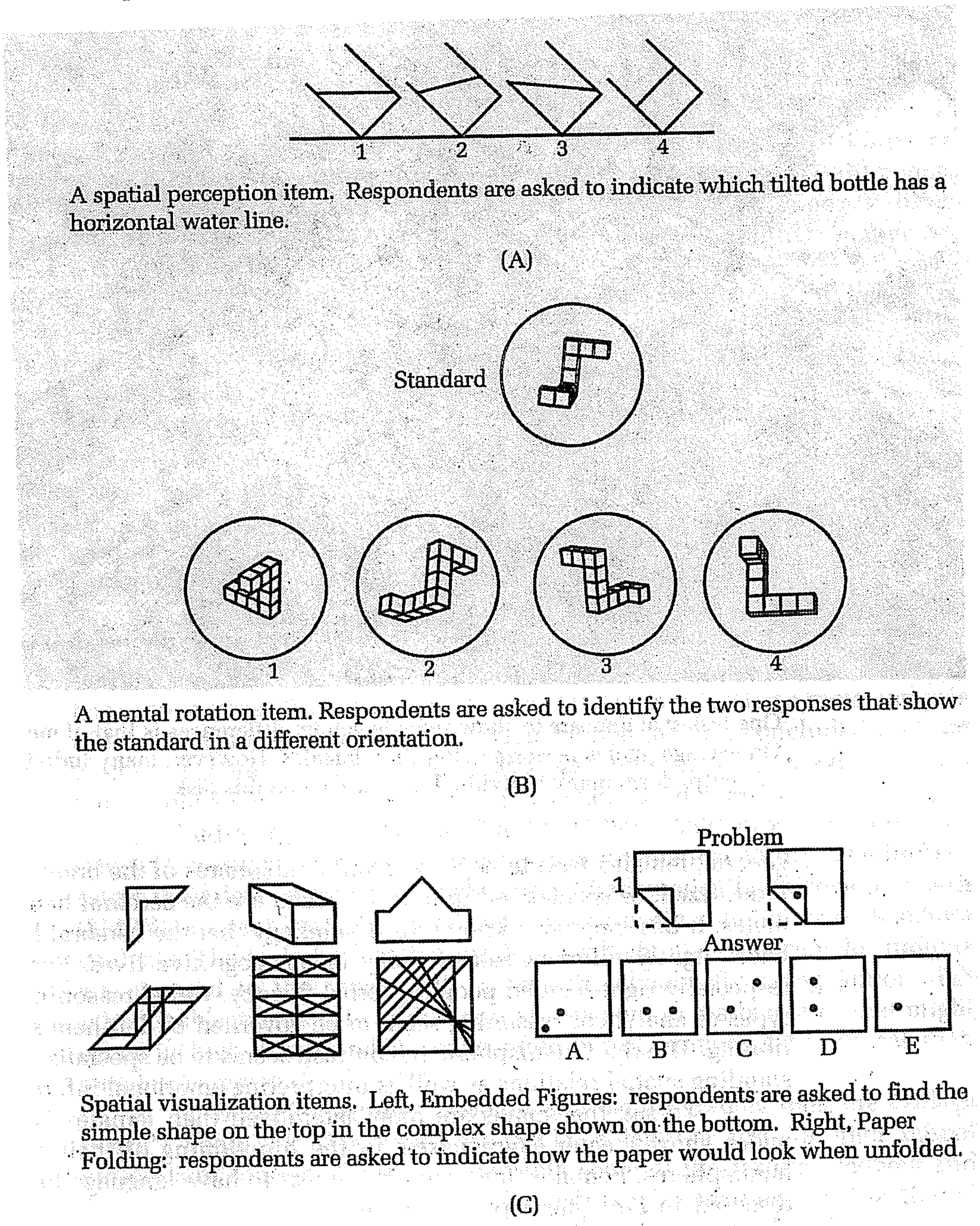
To summarise, spatial abilities are cognitive functions that enable people to deal effectively with spatial relations, visual spatial tasks and orientation of objects in space. Contemporary city inhabitants' spatial abilities may have evolved with the new living and cultural conditions, and the development of new spatially based abilities may have increased the right hemisphere specialisation for handling information with spatial content. The spatial abilities are a part of the fluid intelligence, which is usually assessed with tests of general reasoning ability and with math.

Spatial tests

The spatial tests to be discussed here are mainly from different standardised test batteries. They are all measuring how subjects solve different spatial tasks, and they all consist of a single trial test (performing a task without any study phase or practice). The concept of spatial ability or spatial cognition is vague and there are several approaches to how spatial ability should be defined or classified. To define spatial ability/abilities one must initially determine whether it is a unitary concept or involves a number of diverse components. The problem with using the factor analytical approach to group and define spatial abilities is that it does not necessarily produce converging definitions. The lack of a universally accepted definition of spatial ability may be due to the large variety of tests used in the psychometric studies or the lack of replicability of the factor structures found when several tests are used (Voyer, Voyer and Bryden, 1995). According to Voyer, Voyer and Bryden's (1995) definition of different spatial abilities in terms of tests, each test is considered as an operational definition of one specific component of spatial ability. When trying to separate out different aspects of spatial cognition in terms of mental processes the differences or the mismatch between tests and processes becomes a problem. In solving the tasks in a spatial test different processes could be used and due to that, when defining spatial ability as different processes, solving the tasks in one test could include the use of several processes.

As a result of a meta-analysis of studies made between 1974 and 1982, Linn and Petersen (1985) made a classification of spatial tests into three distinct categories and labelled these categories spatial perception, spatial visualisation and mental rotation (see figure 1). *Spatial perception* was defined as the ability to determine spatial relations despite distracting information and *spatial visualisation* as the ability to manipulate complex spatial information when several stages are needed to produce the correct solution (Linn and Petersen, 1985). *Mental rotation* was defined by Linn & Petersen (1985) as the ability to rotate, in imagination, quickly and accurately two- or three-dimensional figures. A slightly different definition made by Kolb and Whishaw (1990) state that mental rotation is the ability to adopt novel perspectives, to see the other side of things. Furthermore, Kolb and Whishaw (1990) divide the mental rotation aspect of the cognitive space into two categories, visualisation and orientation, according to neurological representations. *Visualisation* is the ability to manipulate or rotate two- or three-dimensional pictorially presented stimulus objects. This visualisation aspect of mental rotation is the same as Linn and Petersen's (1985) overall definition of mental rotation. *Orientation* is the ability to remain unconfused by the changing orientation in which spatial configuration may be presented (Kolb and Whishaw, 1990). The categorisation made by Linn and Petersen (1985) were tested by Voyer, Voyer and Bryden (1995) in an other meta-analysis of studies made between 1974 and 1993. These results showed that the groping of spatial tests made by Linn and Petersen (1985) is replicated reasonable well (Voyer, Voyer and Bryden, 1995). Some spatial tests from different batteries of tests and from different factor analyses, are going to be presented here according to these three categories.

Mental rotation: In Linn and Petersen's (1985) analysis the mental rotation grouping included the Spatial Relations subset of the Primary Mental Abilities Test (PMA-SR) (Thurstone, 1958) and the Cards Rotation Test (CRT) (Ekstrom, French & Harman, 1976), in which participants are required to perform a mental rotation of two-dimensional objects.



A spatial perception item. Respondents are asked to indicate which tilted bottle has a horizontal water line.

(A)

A mental rotation item. Respondents are asked to identify the two responses that show the standard in a different orientation.

(B)

Spatial visualization items. Left, Embedded Figures: respondents are asked to find the simple shape on the top in the complex shape shown on the bottom. Right, Paper Folding: respondents are asked to indicate how the paper would look when unfolded.

(C)

Figure 1. Examples of the three kinds of spatial-ability tasks identified by Linn and Petersen (1985)

In this mental rotation category Linn and Petersen (1985) also included the Mental Rotations Test (MRT) (Vandenber & Kuse, 1978), which is a paper-and-pencil version of the Shepard and Metzler (1971) mental rotation task, using three-dimensional objects. Further this category includes what Voyer, Voyer and Bryden (1995) has termed Generic Mental Rotation Tasks, which include any variant of the Shepard and Metzler(1971) chronometric task, either presented on slides or on a computer screen.

Two spatial tests from the Dureman-Sälde test battery (Psykologiförlaget, 1971), which were not included neither in the meta-analysis made by Linn and Petersen (1985) nor in the meta-analysis made by Voyer, Voyer and Bryden (1995), also loaded high on a mental rotation factor (Dahlbäck, Höök and Sjölinder, 1996). These tests, rotation of images and left or right hand identification, both involve mental rotation of two-dimensional figures. In the rotation of images the subject should choose, by turning the images in their head, the images that were identical with the image in the task. The hand identification task consisted of pictures of hands turned in different ways where the subject should answer if it was a picture of a right hand or a picture of a left hand.

Mental rotation tests:

TEST	NAME	REFERENCE	DESCRIPTION
Spatial Relations subset of the Primary Mental Abilities Test	PMA-SR	Thurstone, 1958	Participants are required to perform a mental rotation of two-dimensional objects
Cards Rotation Test	CRT	Ekstrom, French & Harman, 1976	Participants are required to perform a mental rotation of two-dimensional objects
The Mental Rotations Test	MRT	Vandenber & Kuse, 1978	A paper-and-pencil version of the Shepard and Metzler (1971) mental rotation task, using three-dimensional objects
Generic Mental Rotation Tasks		Voyer, Voyer and Bryden, 1995	Including any variant of the Shepard and Metzler (1971) chronometric task, presented on slides or on a computer screen
Rotation of images		Dureman-Sälde test battery, Psykologiförlaget, 1971	The subject should choose, by turning the images in their head, the images that were identical with the image in the task
Left or right hand identification		Dureman-Sälde test battery, Psykologiförlaget, 1971	Pictures of hands turned in different ways where the subject should answer if it was a picture of a right hand or a picture of a left hand

Spatial perception: The tests included in Linn and Petersen's (1985) spatial perception category were the rod-and-frame test (RFT) (Witkin & Asch, 1948), which requires subjects to adjust a rod to the vertical, despite the distracting information provided by the tilted frame. Further this category included the Water Level Test (WLT) (Piaget & Inhelder, 1956), in which participants are required to indicate the orientation of liquid in a tilted container.

Spatial perception tests:

TEST	NAME	REFERENCE	DESCRIPTION
The rod-and-frame test	RFT	Witkin & Asch, 1948	This test requires subjects to adjust a rod to the vertical, despite the distracting information provided by the tilted frame
The Water Level Test	WLT	Piaget & Inhelder, 1956	Participants are required to indicate the orientation of liquid in a tilted container

Spatial visualisation: In the spatial visualisation category Linn and Petersen (1985) included the Paper Form board (PFB) (Likert & Quasha, 1941) in which participants must decide which of five, two-dimensional line drawings of shapes can be made out of a set of fragmented parts. In the Identical Blocks Test (IBT) (Stafford, 1961) participants must indicate which block, among a number of alternatives, that is the same as a standard. This given a variety of cues, in letters and numbers on the faces of the blocks. This spatial visualisation category is a catchall grouping (i.e. tests that does not fit anywhere else) and these two tests (PFB and IBT) have an important mental rotation component. Therefore they might be better classified in the mental rotation category (Voyer, Voyer & Bryden, 1995). Results with gender differences in the PFB could also indicate that this test includes a mental rotation component (Stumpf & Eliot, 1995).

The spatial visualisation grouping is, according to Linn and Petersen (1985), also incorporated the various adult's and children's version of the Embedded Figures Test (EFT and CEFT) (Witkin, 1950) as well as the Hidden Figures Test (HFT) (Ekstrom, French & Harman, 1976). In these tests participants must find a simple figure embedded within a complex pattern. The Embedded Figures Test also show some characteristics that are included in the definition of spatial perception. This test have at least one component of the ability to determine spatial relations despite distracting information, though it consists of finding specific figures in larger patterns or figures. Another test in the spatial visualisation group is the Paper Folding (PF)(Ekstrom, French & Harman, 1976) In this test the participants are required to indicate which one among four unfolded pieces of paper is the same as a folded model in which holes were punched.

The Block Design (BD) subset of the Wechsler Adult Intelligence Scale, the Wechsler Adult Intelligence Scale-Revised, and the Wechsler Intelligence Scale for Children (Wechsler, 1946, 1949, 1955, 1974, 1981) is a spatial test in which participants must reconstruct a shape using three-dimensional blocks. In a factor analytical study made by Dahlbäck, Höök and Sjölinder (1996) the Block Design loaded high on a factor different from the other spatial tests (rotation of images and the left or right hand identification). The test that clustered with the Block Design was Figure Drawing which is a test that also demands the ability to manipulate complex information where several stages are needed to produce the correct solution. This result fit well with the categorisation made by Linn and Petersen (1985).

This category also included the Differential Aptitude Test - Spatial Relations Subset (DAT-SR) (Bennett, Seasharo & Wesman, 1947), in which participants are required to indicate what an unfolded shape would look like when folded. Kovac and Rensselaer (1989) made an analysis of how user-processing strategies were related to different spatial ability tests and found that DAT showed the poorest performance with numerous invalid user-processing strategies.

Spatial visualisation tests:

TEST	NAME	REFERENCE	DESCRIPTION
Paper Form board	PFB	Likert & Quasha, 1941	Participants must decide which of five, two-dimensional line drawings of shapes can be made out of a set of fragmented parts
DAT Spatial Relations Subset	DAT-SR	Bennett, Seasharo & Wesman, 1947	The subjects are required to indicate what an unfolded shape would look like when folded
Identical Blocks Test	IBT	Stafford, 1961	Participants must indicate which block, among a number of alternatives, that is the same as a standard given a variety of cues (letters and numbers on the faces of the blocks)
The Block Design subset of the Wechsler Adult Intelligence Scale, the Wechsler Adult Intelligence Scale-Revised, and the Wechsler Intelligence Scale for Children	BD	Wechsler, 1946, 1949, 1955, 1974, 1981	Participants must reconstruct a shape using three-dimensional blocks
Paper Folding	PF	Ekstrom, French & Harman, 1976	Participants are required to indicate which one, among four unfolded pieces of paper, that is the same as a folded model in which holes were punched
Various adult and children's version of the Embedded Figures Test	EFT and CEFT	Witkin, 1950	The subjects must find a simple figure embedded within a complex pattern
Hidden Figures Test	HFT	Ekstrom, French & Harman, 1976	The subjects must find a simple figure embedded within a complex pattern

In summary; to make some conclusions about measuring the spatial ability one must first define if it is a unitary concept or involves a number of diverse components. A problem with the factor analytical approach is, that in order to group and define spatial abilities, it does not necessarily produce converging definitions. When trying to separate different aspects of spatial cognition in terms of underlying mental processes, there is a possibility that solving the tasks in one test could include the use of several processes. However spatial tests can be clustered into three categories: spatial perception, spatial visualisation and mental rotation. Mental rotation defined as the ability to rotate, in imagination, quickly and accurately two- or three-dimensional figures. Spatial perception as the ability to determine spatial relations despite distracting information. Finally spatial visualisation, defined as the ability to manipulate complex spatial information when several stages are needed to produce the correct solution.

Acquisition of spatial knowledge and spatial memory

Examples of spatial memory include remembering where a specific article is in the newspaper, finding items around the house and remembering where buildings are geographically located (Pezdec, 1983). Historically, models of spatial representation have taken one or the other of two opposing forms. One proposal is that spatial representations are map-like and preserve Euclidean properties of the world. The other view is that spatial representations are abstract conceptual representations that may or may not preserve Euclidean properties of the world (McNamara, 1991).

According to *hierarchical theories*, spatial memories contain nested levels of detail. The structure of these memories can be expressed in graph-theoretic trees, such that global and local properties of an environment are represented at different levels of a tree. Region membership is an important global property of many spatial environments, where regions can be defined by physical boundaries (walls between rooms), perceptual boundaries (lines on a map), or subjective distinctions (uptown Vs downtown). Locations in the same region of a spatial layout prime each other more than locations in different regions, even when Euclidean distance is held constant and objects in the same region of a spatial layout are closer in subject's memories than objects in different regions (McNamara, 1991). Environments, for instance, are organised hierarchically in memory with features that are larger or functionally more significant having priority over those that are less so (Taylor & Tversky, 1996). *Non hierarchical theories* constitute the second class of theories. The prototypical example of a non hierarchical theory is the mental image, in which spatial relations are represented holistically. They lack hierarchical structure and the important claim is that spatial representations do not contain nested levels of detail or separate codes for global and local properties (McNamara, 1991).

In the acquisition of spatial knowledge a distinction between *primary* and *secondary* spatial learning was made by Presson and Halzerigg in 1984. The definition of primary spatial learning, according to Schacter and Nadel (1991), refer to spatial memories that are based on direct experience of an environment whereas spatial memories that are acquired symbolically is refereed to as secondary. Memory for spatial locations that are learned through real-world navigational experience (primary spatial knowledge) differs from memory for locations that are acquired from maps (secondary spatial knowledge). The best documented difference between the two is that primary spatial memory is largely independent of various orientation manipulations at the time of retrieval. These on the other hand have a significant effect on secondary spatial memory (Schacter & Nadel, 1991). Three main types of spatial knowledge that are acquired when people engage in primary spatial learning are: *landmarks* or *reference points, route knowledge* and *configural knowledge*. Landmarks or reference points refer to salient, distinctive features of a spatial layout that are used as a basis for making judgements about various aspects of the layout. Route knowledge refers to the spatio-temporal relations between specific environmental features. Finally configural knowledge refers to map like structures that include information about the interrelationships among locations. Some developmental evidence suggests that encoding and retrieval of landmark knowledge may precede other sorts of spatial knowledge (e.g. Schacter & Nadel, 1991). In the everyday physical environment paths (order of landmarks) are learned before the locations of landmarks and the acquisition of these spatial properties may be very fast. Once acquired there may be marginal forgetting of these memory representations, at least for retention intervals that people often experience in the town in which they live (Gärling et al., 1981).

Some manipulations of encoding instructions, preceding retention of maps, will affect the retention rate. Two different encoding conditions, in this case, are semantic encoding (semantic relatedness of features, cluster to gain meaningful units) Vs spatial encoding (cluster map features by their spatial proximity, cluster by distance). Semantic encoding instructions yields higher retention rates over instructions to encode spatially, and females profit more from the directive to encode map features semantically (Schwartz & Phillippe, 1991). Further, there are fundamental differences in the processes underlying memory for the locations of objects and words. Item location memory are greater for objects than for words (Pezdek, 1983; Puglisi et al., 1985) and the spatial location is encoded as an integral aspect of the memory representation of objects, but not with the memory representation of words (Schacter & Nadel, 1991).

Hasher and Zacks (1979) suggested that memory for spatial layouts involved automatic processing and therefore would not decline with age. Other findings show that spatial encoding is not automatic and draws upon effortful processes (Lipman & Caplan, 1992; Naveh-Benjamin, 1987; Naveh-Benjamin, 1988; Pezdek, 1983; Puglisi et al., 1985; Salthouse et al., 1989;) or that the distinction of Hasher and Zacks (1979) between automatic and effortful processes is not an adequate distinction for understanding spatial memory (Puglisi et al., 1985). One explanation to these contradicting results concerning the distinction between automatic and effortfull processing of spatial memories could be that there is a dependence on the kind of spatial memory test that is assessed (Schacter & Nadel, 1991).

To summarise this part on acquisition of spatial knowledge. There are two opposite views on the organisation of spatial memories. According to hierarchical theories, spatial memories contain nested levels of detail; and region membership is an important property. Environments, for instance, are organised hierarchically in memory. According to non hierarchical theories spatial relations are represented holistically, as for instance in a mental image.

The acquisition of spatial knowledge can be either primary or secondary. Primary spatial knowledge is acquired by navigation in the real world, and secondary spatial knowledge is acquired, for instance, through maps. When one is engaged in primary spatial learning there are three main types of spatial knowledge that is acquired: landmarks or reference points, route knowledge and configural knowledge. Encoding and retrieval of landmark knowledge precede other sorts of spatial knowledge. When acquiring information of an environment through a map, semantic encoding instructions yields higher retention rates over instructions to encode spatially.

Spatial descriptions

The description of an environment is dependent on the viewer's perspective and of the characteristics of the space to be described. One way of describe a spatial layout is to take the reader on a mental tour or *route* through the environment. A *survey* description on the other hand take a perspective from above and describe the locations of landmarks relative to one another in canonical direction terms: north, south, east, and west (Tversky 1991). A third way of describing an environment is to take the listeners on a *"gaze tour"*. Speakers adopt a fixed point of view from outside the room and describe locations relative to each other in terms of "in front of" and "to the left" with respect to the outside viewpoint. Gaze tours typically locate objects relative to other objects from a fixed outside point of view, so it uses a deictic or viewer-centred frame of reference. In contrast, a route tour typically uses an intrinsic or object-centred frame of reference. A spatial description analogous to a survey viewpoint would use an extrinsic or environment-centred frame of reference (Taylor & Tversky, 1996). Each of the three primary styles of describing environments: gaze, route, and survey reflects a natural way of experiencing an environment and the choice of description perspective depends on the characteristics of the environment and on how an environment has been experienced. In describing networks or rooms that can be seen from one viewpoint, people prefer gaze descriptions. In describing their apartments, people prefer route descriptions (Linde & Labov, 1975; Taylor & Tversky, 1996). In describing their town, people typically uses a survey consistently, or else use a survey for the larger features and a route for the buildings. There are several differences among these environments, any of which may contribute to choice of perspective (Taylor & Tversky, 1996).

A series of studies made by Taylor and Tversky (1992a; 1992b) pointed out that readers form the same spatial mental models capturing the spatial relations between landmarks from both survey and route descriptions, as well as from maps (Taylor and Tversky,1992a). The maps that the subjects in the "description condition" constructed, solely from descriptions, were quite similar to the maps the subjects in the "map condition" study, and the descriptions these subjects constructed from the maps were quite similar to those that the subjects in the "description condition" read (Taylor and Tversky, 1992b). Other findings suggest that there are differences between visual and verbal representations of the same space (Stock et. al., 1995). An explanation to the difference, in favour of maps, could be that maps are represented as images and that these images, relative to representations derived from verbal descriptions, free resources in working memory that can then be used to process text (Stock et. al., 1995). When using a text to describe a spatial environment, landmarks used in the text as reference points for describing the locations of some other landmarks are represented more accurately than the location of other landmarks. This is independent of the perspective from which the text was written (route or survey) or whether or not a map was present at learning (Ferguson & Hegarty, 1994).

To summarise, an environment can be described from three different perspectives. A survey description take a perspective from above, a route description take the reader on mental tour through the environment and a gaze tour locate objects relative to other objects from a fixed outside point of view. Each of these different ways of describing an environment reflects a natural way of experiencing an environment and the choice of description perspective depends on the characteristics of the environment.

Summary

Spatial abilities are cognitive functions that enable people to deal effectively with spatial relations, visual spatial tasks and orientation of objects in space. Contemporary city inhabitants' spatial abilities may have evolved with the new living and cultural conditions, and the development of new spatially based abilities may have increased the right hemisphere specialisation for handling information with spatial content. The spatial abilities are a part of the fluid intelligence, which is usually assessed with tests of general reasoning ability and with math.

When measuring the spatial ability, first one must define if it is a unitary concept or involves a number of diverse components. A problem with the factor analytical approach is, that in order to group and define spatial abilities, it does not necessarily produce converging definitions. When trying to separate different aspects of spatial cognition in terms of underlying mental processes, there is a possibility that solving the tasks in one test could include the use of several processes. However spatial tests can be clustered into three categories: spatial perception, spatial visualisation and mental rotation. Mental rotation defined as the ability to rotate, in imagination, quickly and accurately two- or three-dimensional figures. Spatial perception as the ability to determine spatial relations despite distracting information. Finally spatial visualisation, defined as the ability to manipulate complex spatial information when several stages are needed to produce the correct solution.

There are two opposite views on the organisation of spatial memories. According to hierarchical theories, spatial memories contain nested levels of detail; and region membership is an important property. Environments, for instance, are organised hierarchically in memory. According to non hierarchical theories spatial relations are represented holistically, as for instance in a mental image. The acquisition of spatial knowledge can be either primary or secondary. Primary spatial knowledge is acquired by navigation in the real world, and secondary spatial knowledge is acquired, for instance, through maps. When one is engaged in primary spatial learning there are three main types of spatial knowledge that is acquired: landmarks or reference points, route knowledge and configural knowledge. Encoding and retrieval of landmark knowledge precede other sorts of spatial knowledge. When acquiring information of an environment through a map, semantic encoding instructions yields higher retention rates over instructions to encode spatially.

An environment can be described from three different perspectives. A survey description take a perspective from above, a route description take the reader on mental tour through the environment and a gaze tour locate objects relative to other objects from a fixed outside point of view. Each of these different ways of describing an environment reflects a natural way of experiencing an environment and the choice of description perspective depends on the characteristics of the environment.

ACKNOWLEDGEMENTS

Thanks to all participants in the PERSONA group for feedback and to HSFR (The Swedish Council for Research in the Humanities and Social Sciences) for additional founding for this work.

References

Ardila, A. (1993). "Historical evolution of spatial abilities." *Behavioural Neurology Vol. 6:* 83-87.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1947). *Differential Aptitude Test.* New York, Psychological Corporation.

Dahlbäck, N., Höök, K., & Sjölinder, M. (1996). Spatial cognition in the mind and in the world - the case of hypermedia navigation. *The Eighteenth Annual Conference of the Cognitive Sciences Society,* University of California, San Diego, July 12-15, 1996, Lawrence Erlbaum Associates, Inc.

Collins Dictionary of the English Language, Second Edition. (1986). Glasgow. William Collins Sons & Co. Ltd.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests.* Princetown, NJ, Educational Testing Service.

Ferguson, E. L., & Hegarty, M. (1994). "Properties of cognitive maps constructed from texts." *Memory and Cognition Vol. 22(No. 4):* 455-473.

Gärling, T., Böök, A., Lindberg, E., & Nilsson, T. (1981). "Memory for the spatial layout of the everyday physical environment: Factors affecting rate of acquisition." *Journal of Environmental Psychology Vol. 1:* 263-277.

Hasher, L., & Zacks, R. T. (1979). "Automatic and effortful processes in memory." *Journal of Experimental Psychology: General Vol. 108:* 356-388.

Kolb, B., & Whishaw, I. Q. (1990). *Fundamentals of Human Neuropsychology.* New York, W. H. Freeman and Company.

Kovac, R. J., & Rensselaer, B. (1989). "The validation of selected spatial ability tests via correlational assessment and analysis of user-processing strategy." *Educational Research Quarterly Vol. 13(No. 2):* 26-34.

Likert, R., & Quasha, W. H. (1941). *The revised Minnesota Paper Form Board.* New York, Psychological Corporation.

Linde, C & Labov W. (1975) "Spatial networks as a site for the study of language and thought. " *Language Vol. 51:* 924-939.

Linn, M. C., & Petersen, A. C. (1985). "Emergence and characterisation of gender differences in spatial abilities: A meta-analysis." *Child Development Vol. 56:* 1479-1498.

Lipman, P. D., & Caplan, L. J. (1992). "Adult Age Differences in Memory for Routes: Effects of Instructions and Spatial Diagram." *Psychology and Aging Vol. 7(No. 3):* 435-442.

McNamara, T. P. (1991). "Memory's View of Space." *The Psychology of Learning and Motivation Vol. 27:* 147-186.

Naveh-Benjamin, M. (1987). "Coding of spatial location information: An automatic process?" *Journal of Experimental Psychology: Learning, Memory and Cognition Vol. 13:* 595-605.

Naveh-Benjamin, M. (1988). "Recognition memory of spatial location information: Another failure to support automaticity." *Memory & Cognition Vol. 16(No. 5):* 437-445.

Oxford Advanced Learner's Dictionary of Current English, Encyclopedic Edition. (1992). Oxford. Oxford University Press.

Pezdek, K. (1983). "Memory for Items and Their Spatial Locations by Young and Elderly Adults." *Developmental Psychology Vol.19(No. 6):* 895-900.

Piaget, J., & Inhelder, B., (1956). *The child's conception of space.* London, Routledge & Kegan Paul.

Psykoloigiförlaget (1971). *Manual for the Dureman-Sälde tests.* Stockholm, Psykologiförlaget AB.

Puglisi, J. T., Park, D. C., Smith, A. D, & Hill, G. W. (1985). "Memory for two types of spatial location: Effects of instructions, age, and format." *American Journal of psychology Vol.98(No. 1):* 101-118.

Reber, A. S. (1985) *The Penguin Dictionary of Psychology.* England, Clays Ltd.

Rolfhus, E. L., & Ackerman, P. L. (1996). "Self-Report Knowledge: At the Cross-roads of Ability, Interest, and Personality." *Journal of Educational Psychology Vol. 88(No. 1):* 174-188.

Salthouse, T. A., Mitchell, D. R. D., & Palmon, R. (1989). "Memory and Age Differences in Spatial Manipulation Ability." *Psychology and Aging Vol. 4(No. 4):* 480-486.

Schacter, D. L., & Nadel, L. (1991). Varieties of Spatial Memory: A Problem for Cognitive Neuroscience. *Perspectives on Cognitive Neuroscience.* R. G. Lister, & Weingartner, H. J. New York, Oxford, Oxford University Press.

Schwartz, N. H., & Phillippe, A. E. (1991). "Individual Differences in the Retention of Maps." *Contemporary Educational Psychology Vol. 16:* 171-182.

Shepard, R. N., & Metzler, J. (1971). "Mental rotation of three dimensional objects." *Child Development Vol. 171:* 701-703.

Stafford, R. E. (1961). "Sex differences in spatial visualisation as evidence of sex-linked inheritance." *Perceptual and Motor Skills Vol. 13:* 428.

Stock, W. A., Kulhavy, R. W., Peterson, S. E., Hancock, T. E., & Verdi, M. P. (1995). "Mental Representations of Maps and Verbal Descriptions: Evidence They May Affect Text Memory Differently." *Contemporary Educational Psychology Vol. 20:* 237-256.

Stumpf, H., & Eliot, J. (1995). "Gender-related differences in spatial ability and the k factor of general spatial ability in a population of academically talented students." *Person. individ. Diff. Vol. 19:* 33-45.

Taylor, H. A., & Tversky, B. (1992a). "Spatial Mental Models Derived from Survey and Route Descriptions." *Journal of Memory and Language Vol. 31:* 261-292.

Taylor, H. A., & Tversky, B. (1992b). "Descriptions and depictions of environments." *Memory and Cognition Vol. 20(No. 5):* 483-496.

Taylor, H. A., & Tversky, B. (1996). "Perspectives in Spatial Descriptions." *Journal of Memory and Language Vol. 35:* 371-391.

Thurstone, T. G. (1958). *Manual for the SRA Primary Mental Abilities.* Chicago, Science Research Associates.

Tversky, B. (1991). "Spatial Mental Models." *The Psychology of Learning and Motivation Vol. 27.*

Vandenberg, S. G., & Kuse, A. R. (1978). "Mental rotation, a group test of three-dimensional spatial visualisation." *Perceptual and Motor Skills Vol. 47:* 599-604.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). "Magnitude of Sex Differences in Spatial abilities: A Meta-Analysis and Consideration of Critical Variables." *Psychological Bulletin Vol. 117(No. 2):* 250-270.

Wechsler, D. (1946). *The Wechsler-Bellevue intelligence Scale, form II.* New York, Psychological Corporation.

Wechsler, D. (1949). *The Wechsler Intelligence Scale for Children.* New York, Psychological Corporation.

Wechsler, D. (1955). *Manual for the Wechsler Adult intelligence Scale.* New York, Psychological Corporation.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children-Revised.* New York, Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* New York, Psychological Corporation.

Witkin, H. A., & Asch, S. E. (1948). "Studies in space orientation. IV. Further experiments on perception of the upright with displaced visual fields." *Journal of Experimental Psychology Vol. 38:* 762-782.

Witkin, H. A. (1950). "Individual differences in ease of perception of embedded figures." *Journal of Personality Vol. 19:* 1-15.

Chapter 5

Individual differences in spatial cognition and hypermedia navigation

Marie Sjölinder
SICS

One mental ability crucial for navigation, both in the real world and in a virtual environment, is spatial cognition. This review will discuss individual differences in spatial cognition and wayfinding; and how these differences could effect navigation in a virtual environment. Other human aspects, as cognitive strategies and personality factors, related to navigation and the use of computers will also be discussed. .

Individual differences in spatial cognition and hypermedia navigation

Marie Sjölinder
SICS Box 1263 S-164 28 Kista Sweden[1]
marie@sics.se
http://www.sics.se/humle/

One mental ability crucial for navigation, both in the real world and in a virtual environment, is spatial cognition. This review will discuss individual differences in spatial cognition and wayfinding; and how these differences could effect navigation in a virtual environment. Other human aspects, as cognitive strategies and personality factors, related to navigation and the use of computers will also be discussed.

Introduction

Different people find their way to their destinations in different ways. They use different representations of the environment, and choose different landmarks or reference points. There are for instance sex differences in some aspects of spatial ability and age differences in representing the environment and in wayfinding. Another issue that effects navigation tasks is the use of different cognitive strategies. Different cognitive styles are related to the ability to learn conceptual material and to the way a person structures and processes information. We shall here investigate a number of factors that to varying degrees are connected with spatial cognition and our wayfinding ability. These factors are: development, sex, age, cognitive strategies, practice, cognitive style and personality.

We will also investigate the connection between these individual differences and hypermedia navigation; and to what extent they influence wayfinding and navigation in a virtual environment. For instance, individuals with a low spatial ability or low navigational ability could, if they have problems with finding their way around in a hypermedia environment, benefit from additional non-spatial information. Different ways of learning and organising a material also may effect the way a person learns to use a computer system. The purpose of this review is to provide input to the design of computer interfaces and navigational instruments, especially with respect to individual differences and how to provide individuals with low navigation ability additional help.

Spatial cognition and development

An important aspect when talking about spatial ability or spatial cognition is to decide if it is a unitary concept or not. One way of clarifying the issues involved here, is to look at developmental aspects of spatial cognition. One way of arguing for a multidimensional view of spatial cognition is that different spatial tasks have different ages-of-onset (Rosser, 1994), that is that the child has to have reached a certain age in order to succeed in performing a particular task. One way of distinguishing different spatial behaviours according to developmental aspects is to distinguish between position, cue and place responses. Position responses are egocentric, in the sense that they are made with respect to the persons own body. Cue responses are movements that are directed by a particular cue, for example walking to or from an object. Both these responses are evolved quite early in life. Place responses takes a person to a particular location or object, which even may be hidden from the view. Usually the relational properties of surrounding cues guide movement. These place responses evolve as the neocortex develops (Kolb and Whishaw, 1990). Another approach or a complementary approach to the development of

[1] Also with the Department of Psychology, Stockholm University, and the Department of Computer and Information Science, Linköping University

spatial cognition is based on activity theory and emphasises the goal of an activity and the socio-cultural context. The development of spatial ability is dependent on socio-cultural concepts, like, for instance, the convention of describing the environment in terms of a map. The individuals' spatial ability develops in the light of these conventions, while being a part in performing activities with other member of the society. The younger individuals get socialised in to a certain way of for instance describing routes or in the comprehension of understanding maps and spatial descriptions (Gauvain, 1993).

To summarise, there is some developmental evidence for a multidimensional view of spatial cognition, with different ages-of-onset for different aspects of the spatial ability. Differences in the spatial ability could also, as a complementary approach, be explained in terms of socio-cultural differences.

Sex differences

Sex differences arise on some types of spatial ability but not others. The largest sex differences in favour of males are found only on different measures of mental rotation and smaller differences are found on measures of spatial perception (Linn & Petersen, 1985). There are small differences or advantages for females on visual memory tasks (Stumpf and Eliot, 1995). When sex differences are found, they can be detected across the lifespan (Linn & Petersen, 1985) but the age of emergence of sex differences depends on the test used (Voyer, Voyer & Bryden, 1995).

Several studies show that at least some of the gender differences could be related to socio-cultural or educational factors (Baenninger & Newcombe, 1989; Gittler & Vitouch, 1994; Hamilton, 1995; Lawton, 1994; Sharps, Price & Williams, 1994). When there are no explicit spatial test instructions (subjects are not told that it is a spatial task, they are just told what to do) there are no sex differences in performing mental image rotation tasks (Sharps, Price & Williams, 1994). Women are culturally expected to exhibit diminished spatial cognitive performance relative that of men. Men constantly rate their spatial abilities as superior to those of women. On the contrary, women are repeatedly subjected to the expectation that they cannot, or may not, excel in spatial behaviours generally. The negative feelings of women toward spatial cognitive activities that may violate feminine self-concept, could be very damaging to performance on spatial tasks such as mental image rotation, as a result of diminished motivation to perform in what is perceived as a gender-inappropriate activity (Sharps, Price and Williams, 1994). Consistent with Sharps, Price and Williams' (1994) findings Hamilton (1995) reported results that indicated that in a 3-D mental rotation task where a substantial sex difference occurred, self-perceived gender trait possession adds significantly to the overall explanation of performance.

Two meta-analyses made by Baenninger & Newcombe (1989) provide evidence for the hypothesis that spatial experience is related to good spatial test performance. Prior participation in spatial activities (sports, games, real life settings) is correlated with higher spatial test scores, although the magnitude of this effect is very small, possibly due to the unreliability of retrospective measures of activity participation. Another aspect of participating in spatial activities is that this participation could be due to a higher level of spatial ability. By contrast spatial training - doing the same test several times - appears to give a fairly strongly, but possible narrow, improvement on spatial test scores. This is true for both sexes and does not differently improve the scores of males and females. The hypothesis drawn from the same results would then predict that females would need more training than males to reach the same level of spatial performance (Baenninger & Newcombe, 1989).

According to Lawton (1994) there are gender differences in self-reported use of different wayfinding strategies. Women were more likely to report using a rote strategy (attending to instructions on how to get from place to place), whereas men were more likely to report using an orientation strategy (maintaining a sense of their own position in relation to environmental reference points). The use of maps are differently associated with the two way-finding strategies: use of published road maps are associated with the orientation strategy, whereas use of hand-drawn maps are linked to the route strategy (Lawton, 1994). Women also report higher levels of spatial anxiety, or anxiety about environmental navigation, than men do. Spatial anxiety is negatively related to the orientation way-finding strategy. One possible explanation for this relationship is that individuals who do not maintain a sense of their own position with respect to the overall environment are more likely to become confused and anxious. Alternatively, anxiety about becoming lost may reduce ability to focus on the cues necessary to maintain geographical orientation in the same way that stress, in form of noise or crowding, reduces memory for spatial locations (Lawton, 1994).

There are a smaller gender differences in orientation strategy in younger subjects, which may be the result of changing gender roles, with more equal opportunity for men and women to engage in way-finding activities, and/or a diminishing influence of the stereotype that men have superior way-finding ability compared to women (Lawton, 1994). Gittler and Vitouch (1994) underline the importance of including an "environmental view" to explain a family correlation structure which they observed in their results. Members of the younger generation scored significantly higher on the Three-dimensional Cube Test than their parents did. An interaction illustrates a significant sex difference in favour of males in the older generation while there is none in the younger generation. These differences in performance in the parental but not in the younger sample are easier explained by an environmental rather than a genetic approach. In modern school, boys and girls may show similar performance because education is similar (Gittler & Vitouch, 1994). It is important to point out in this context, that also for the highly reliable gender differences, the percentage of the variance accounted for by gender is only between 1% and 5% (Hyde, 1981).

To summarise, the largest sex differences in favour of males could be addressed to the mental rotation aspect of spatial cognition. Recent research indicate that these sex differences could, however, be a consequence of socio- cultural differences, as for instance gender roles.

Age differences

Hasher and Zacks (1979) suggested that spatial memory is an automatic process (not demanding with respect to cognitive functions), and because of this there will be no age differences. Several studies testing this hypothesis from different perspectives has gained conflicting results. Pezdek (1983) examined age differences in the relationship between memory for items and memory for the spatial location of items. The young adults were more accurate on both tests. This was explained in terms of more effective encoding and rehearsal strategies in favour of the younger adults (Pezdek, 1983). Behind these differences in encoding and rehearsal strategies could be limitations in working memory capacity. The concept of working memory, according to Baddley (1986), involves both a processing function and a storage function. Both these components of working memory are important for predicting performance on spatial thinking and language processing (Shah & Miyake, 1996). The decline in working memory capacity in older adults lies in the integration of these two working memory functions. If memory is assessed in the context of other ongoing tasks, then the individual is required to remember the relevant information while also engaged in attempting to process information. These joint demands of storage and processing may present special difficulties for older adults (Salthouse et al., 1989).

One might expect a diagram or another analogical model (some kind of simplified representation of the space) to be particularly helpful for older adults, who have difficulty in constructing mental networks of spatial relations'. However, older adults might be less likely than younger adults to benefit from such interventions (Caplan & Schooler, 1990; Lipman & Caplan, 1992). High performance with analogical models reflects a predominant use of well-constrained transformations. These are often holistic and spatial in nature (Schiano et al., 1989). Caplan and Schooler (1990) found that the performance of older adults were equal to that of younger adults when no analogical model was provided as a learning aid, but the performance of older adults was significant poorer than of younger adults when a model was provided. To explain these results Caplan and Schooler (1990), pointed out that retrieval is easier when conditions are similar to those during encoding (Morris et al., 1977) or, in this case, acquiring information is easier when the source of external information is coherent with the on-going internal processes. Caplan and Schooler (1990) suggested that models and diagrams are effective learning aids only when people are already engaged in the processing of those underlying rules or relationships, such as that involved in learning route or configural information, which are more likely to be true of younger adults (Lipman, 1991). Under such conditions, the aid provides a useful retrieval cue or context for learning (Caplan and Schooler, 1990). In the case of older adults, who are less likely to be engaged in the processing of configural information, there will be a mismatch between the model and their on going spatial processes. In contrast, when people are engaged in learning about discrete objects or features, such as landmarks, which is more likely to be true of older adults (Lipman, 1991), learning should be best without such aids Caplan and Schooler (1990).

Lipman and Caplan (1992) tested if unambiguous instructions could compensate for the decline, due to age, in the effectiveness to use encoding and retrieval strategies. Memory for routes consists of memory for different aspects of a route (memory for landmarks and turns, their sequential ordering, and their configural representation). Two kinds of representation may underlie these four components of route memory (scene representation and layout representation). Scene representation is the representation of discrete feature of the environment (for instance landmarks) and layout representation is the Euclidean survey representations of the route (for instance configural representation). Memory for landmarks may rely most heavily on scene representation and in contrast, configural memory may rely most heavily on layout representation. Memory for landmark sequence and memory for turns may involve both kinds of information. Some of these aspects of route memory may be more susceptible to age differences than others.

Older subjects' free recall of a route include more conspicuous landmarks in non sequential order and more personal or evaluative statements about the scenes or the testing experience. Younger subjects, on the other hand, are more likely to recall landmarks in sequential order. This suggests that although older individuals may be able to encode and remember scene-based aspects of a route, such as landmarks they may have difficulty with the more integrative aspects involved in layout memory (e.g., landmark order, turns or route shape) Lipman and Caplan (1992).

To summarise this part on age differences in wayfinding, route memory may involve both scene and layout representation, which may be differentially sensitive to age. The age differences are found in configural and integrative aspects of spatial memory and could be explained in terms of older adults disposition towards a more personal and less holistic view on the environments; and/or due to limitations in working memory capacity. Older adults may benefit more from explicit verbal instructions to focus attention on the path or route than to be provided with a model or an analogy over the space. Different strategies of using landmarks and turns could also promote the acquisition of other aspects of spatial knowledge.

Strategies behind performance on spatial activities

According to Messick (1994) the focus in learning styles studies should be on organisation and on control of the strategies involved in knowledge acquisition. There are two different major strategies for knowledge acquisition based on Craik and Lockharts (1972) concept of levels of processing. One strategy is holistic, based on an overall meaning, emphasising a conclusion-oriented deep-processing approach to learning, while the other strategy is more serialistic with sequential procedures associated with a description-oriented shallow-processing approach (Messick, 1994).

Magliano et al. (1995) showed the importance of different intentions or goals in the acquisition of spatial knowledge. Subjects in different groups where instructed to attend to different types of spatial information during learning, they were given instructions that either emphasised landmarks, routes, configurations or no specific spatial instructions. No differences in performance attributable to instruction condition occurred for either the landmark recognition or route-sequencing tasks. In an orientation task, subjects given configuration instructions outperformed those receiving landmark instructions. In general, all subjects provided landmark information; subjects given route or configuration instructions provided more route and configuration information. These results suggest that wayfinders are capable of learning a new environment according to a goal, but that learning is constrained by stage-based processes where route knowledge may be dependent upon sufficient landmark knowledge, and configural knowledge may be dependent upon sufficient route knowledge (Magliano et al., 1995). These findings are interesting in relation to the characteristics of older adults wayfinding, which emphasises landmarks (Lipman, 1991), which, according to Magliano et al. (1995), in some sense is the lowest level in the stage-based process of wayfinding.

A possible conclusion to be drawn here, is that there will be no benefits from focus attention on landmarks. The recall of landmarks is at the same level, with or without an emphasis on landmarks in the acquisition phase. On the other hand Craik and Lockhart's (1972) theory on levels of processing is supported by Magliano's et al. (1995) description of a stage-based process. According to this there would be benefits from using a deep-processing approach, going through all the stages (acquisition of landmarks, gaining first route knowledge and then configural knowledge) in this certain order, in the acquisition of spatial knowledge.

Spatial cognition and practice

Practice on spatial tasks and tests can produce substantial improvements in performance, particularly in the speed of response. Practice effects are also greater for simple tests than for more complex tests. This has the implication that the validity of speeded spatial tests could be markedly reduced if subjects have different levels of familiarity with test items. If trying to avoid this problem with using relatively complex, unspeeded spatial tests another problem occurs: these kind of tests tend to be highly correlated with general reasoning or fluid abilities, and due to that provide relatively little information on those abilities that are uniquely spatial (Lohman & Nichols, 1990). Improvement in spatial skills is to a great extent task specific and therefore large improvements on one spatial task does not mean that spatial abilities have been improved significantly. Spatial cognitive skills may be most improvable in subjects who have not much practised performing these skills previously. Practice on a wide variety of tasks should encourage generalisation and the largest changes in spatial ability may be produced not by training in particular skills, but by experiences that permit the gradual accretion of a richer knowledge base (Lohman & Nichols, 1990).

Instructions, based on the characteristics of high performers mental activities, can improve a learners ability to visualise objects from multiview drawings when the design of a computer program allows the learner to se the relationship's between the features on a 3-D wireframe model and its 2-D representation (Duesbury and O'Neal, 1996). McClurg and Chaillé (1987) found that certain computer games enhanced the development of spatial ability measured by a mental rotation test. Their results showed that both boys and girls benefited from the practice but that males use computers and play video games more often than females. However, when girls do play computer games that facilitate the development of spatial ability the effects are as great as when boys play those games (McClurg and Chaillé, 1987). On the other hand, Greenfield and Lohr (1994) found no short-term effects from video game experience on spatial ability measured by a mental paper folding task. However, structural equation modelling did provide strong evidence that video game expertise, developed over the long-term, had a beneficial effect on the spatial skill of mental paper folding (Greenfield and Lohr, 1994).

To summarise the effects on spatial cognition and practice, improvements in performance are task specific. It is difficult to study for how long these effects lasts, because there is a huge amount of variables that are difficult to control and which could effect the subjects performance. The acquisition of a richer knowledge base and the improvement of general reasoning ability will also effect the spatial ability, therefore it is difficult (if the task specific effect lasts) to know if it is really the task specific practise that is lasting, or if it is an over all improvement in the general reasoning ability.

Cognitive style and personality factors

A distinction between intellectual ability and cognitive style is that intellectual ability refer to the context and the level of cognition while cognitive style refer to the manner or mode of cognition. Furthermore, abilities are seen as unipolar (range from none to a great deal with increasing levels) whereas cognitive styles are typically conceived to be bipolar (range from extreme to a contrasting extreme). Abilities are value directional, having more of an ability is better than having less. On the other hand cognitive styles are value-differentiated, each stylistic extreme has adaptive value, but in different circumstances. Abilities are specific to a particular domain of content or function, such as verbal, numerically or spatial ability. A cognitive style, in contrast cuts across domains. Finally, abilities are enabling variables in the sense that they facilitate task performance in specific areas, whereas cognitive styles are organising and controlling variables (Tiedemann,1989). The distinction between learning styles and cognitive styles is, according to Messick (1994), that the focus of learning styles is on the organisation and control of strategies for learning and knowledge acquisition (Messick, 1994), for example different levels of processing (Craik & Lockhart, 1972). Cognitive styles on the other hand focuses on the organisation and control of cognitive processes (Messick, 1994).

One cognitive style that is discussed together with spatial ability is field-dependence Vs field-independence (Schwartz & Phillippe, 1991). A field-independent person is characterised as analytic, self-referent, and impersonal in orientation, whereas the field-dependent person is characterised as global, socially sensitive and interpersonal in orientation (Messick, 1994). Field-independence Vs field dependence is often measured with EFT (Embedded Figures Test), a test where the subjects' task is to find hidden figures in a pattern. Field-independence is an ability closely related to flexibility (Messick, 1994). In a factor analytical study made by Bostic and Tallent-Runnels (1991) field-independence (measured by EFT) clustered together with tolerance for ambiguity (measured by MacDonald Scale of Ambiguity Tolerance) and cognitive complexity (measured by Bieri Repertory Test) under the label of flexibility. It seems that field-independence Vs field-dependence has become a value directional variable in favour of field-independent with its social qualities and analytical traits (Messick, 1994). Cognitive styles influence not only the use of cognitive structures but also their development. As a consequence, cognitive styles have implications for learning and the structuring of knowledge. Field-independent learners appear to be more efficient in the selection and implementation of strategies that co-ordinate information processes. Especially those used in selectively attending to relevant cues, as well as in storing and retrieving information from memory. Field-dependent learners appear to be more responsive to salient cues, whether relevant or irrelevant, and less strategic in orientation (Messick, 1994). Cognitive style has a robust effect on map retention with field-independent learners remembering significant more of a map's features than learners who are field-dependent (Schwartz & Phillippe, 1991).

The cognitive styles dimension of field-dependent Vs field-independent also affect the way a person structures and processes information. This may in turn affect the way a person learns to use a computer system. When learning to use the UNIX operating system field-dependent subjects are less likely to know the command, and more likely than field-independent subjects to ask for help without making any attempt at the task. Field-independent subjects on the other hand are more likely to attempt the task and make errors than ask for help (Coventry, 1989). Field-dependent learners are less likely to impose a meaningful organisation on a field that lacks structure and are less able to learn conceptual material when cues are not available. Field-dependent students fails to take advantage of learner control in selecting options. In the case of hypermedia-based instruction (HBI) field-independent students learn more effectively from HBI than field-dependent students (Weller, 1995).

The Eysenckian personality dimensions of extraversion/introversion correlates with learning style in a way where extraverts are quite clearly activists and pragmatists and introverts reflectors (Furnham, 1992). *Extraversion* also correlates positively with converger learning styles, which is a preference to gather information through abstract conceptualisation. Convergers use deductive reasoning and prefer the application of ideas. They prefer dealing with technical tasks and problems rather than with social and interpersonal issues. Making decisions too quickly and solving the wrong problem has been identified as the weakness of convergers. *Introversion* on the other hand correlates positively with accommodation preferences which is a preference to gather information through concrete experience. The accommodator is quick to involve himself in new situations in a trial-and-error-manner, risk taking is a strength of this style. The term accommodation is so named because these traits are best suited for those situations in which one must adapt oneself to changing immediate circumstances. Individuals with accommodative learning styles are at ease with people but are sometimes seen as impatient and "pushy" (Furnham, 1992).

The introversion-extraversion dimension appears to be an important dimension that is related to many aspects of human-computer interaction. These personality dimensions may be suitable for determining whether an individual will be drawn toward a computer-related career. On the other hand, these personality dimensions are not an effective predictor of performance reflecting programming aptitude or achievement, or of academic achievement in courses based on computer-assisted instructions, (Pocius, 1991). There is a pattern of positive correlations of arts and humanities knowledge (crystallised abilities) with typical intellectual engagement and openness, and a pattern of correlations between math and physical sciences knowledge (fluid abilities) and realistic and investigative interests (Rolfhus and Ackerman, 1996).

Cognitive styles have also been shown to be an important determinant of computer anxiety (measured by CARS, an Computer Anxiety Rating Scale developed by Heinssen at al. and described in Computers in human Behaviour, 1987, 3, 49-59) (Chu and Spires, 1991). Chu and Spires (1991) use Jung's two basic dimensions, perception and judgement, which cognitive styles differs along. *Perception* refers to information acquisition. It is bounded on one end by the sensation-oriented individual and on the other by the intuitive individual. The sensing person tends to focus on what is visible to the senses, the immediate experience and the concrete objects. In contrast , the intuitive person perceives beyond the senses and prefers to ponder the relationships of objects rather than the objects themselves. An intuitive person is at home with theoretical and abstract issues. *Judgements* refers to how problems are resolved. At one extreme is the feeling individual who takes into account individual feelings, considers value to be an important criterion, and is people-oriented. In contrast, the thinking individual tends to be impersonal in the evaluation, relying on logical analysis to guide decision making. The thinking type attempts to generalise from a logical base whereas the feeling type seeks to understand the personalities affected by the decision. The relation between these cognitive styles (sensing Vs intuitive and feeling Vs thinking) and computer anxiety measured by CARS shows that intuitive and thinking individuals exhibiting lower anxiety than their sensing and feeling counterparts (Chu & Spires, 1991). In short, individuals exhibiting lower computer anxiety is at home with theoretical and abstract issues and attempts to generalise from a logical base.

To summarise, intellectual abilities are specific to a particular domain or function whereas a cognitive style cuts across domains. The cognitive style field-independence Vs field-dependence have effect on map retention. Field-independent learners remember more of a map's features than learners who are field-dependent. In the use of computers, field-dependent subjects are less likely to know the command, and more likely than field-independent subjects to ask for help without making any attempt at the task. The personality dimension introversion/extraversion may be suitable for determining whether an individual will be drawn toward a computer-related career. Individuals who are intuitive and thinking exhibit lower computer anxiety. These individuals are familiar with theoretical and abstract issues and attempt to make generalisations from a logical base.

Supporting individual differences in HCI design

Several of the individual differences discussed above refers to real life settings. It is an open question whether, or to what extent, real life navigation is comparable with navigation in a virtual world. Further research on similarities differences between real world navigation and hypermedia navigation will be needed. Another question is to what extent it is possible to make generalisation from studies in individual differences in a real life setting when dealing with a virtual environment. These issues are important when the intentions behind interface design is to take into consideration the individual's social, cognitive and biological presuppositions.

In cognitive neuroscience the function of the "where" cortical processing system directs attention to locations in space. The "what" cortical processing system identifies information. Spatial information detected by the "where" system indexes content information. A principle of interface design which takes spatial indexing into account would state that a spatial code should provide the range of possible behaviours available from each location within the application. The information space must be presented to users from each location-relative point of view as they navigate through the space rather than from an absolute, location-independent point of view as if they are looking down from above. *The principle of natural affordance* states that an interface, at any time, must present both the content of the screen and the range of behaviours that are possible from that point of view. Such behaviours should be indicated in a form that is analogous to peripheral vision by using a dynamic spatial analogue for the information space that changes according to the viewpoint of where the user is in the space (Strong & O'Neill Strong, 1991). A methodology like this will allow the user to understand the applications conceptual model through the behaviour that it affords. In this way users will be able to directly carry out their tasks without having to learn the application and keep it in their mind along with the task they want to perform (Strong & O'Neill Strong, 1991).

Individuals with low spatial ability have difficulties in constructing, or do not use, a mental model of the space. Stanney and Salvendy (1995) tested an interface that would challenge users with low spatial ability to construct a mental model. They developed two human-computer interfaces which were intended to compensate for the inability of low spatial individuals to readily construct visual mental models of a menu systems structure. The information search performance of high and low spatial individuals was compared on the two compensatory interfaces and a third challenge match interface, which challenged individuals to construct a mental model of a hierarchical menu system in order to perform efficiently. The visual mediators were successful in accommodating low spatial individuals, as indicated by the lack of any significant performance differences being detected between the high and low spatial groups on the two compensatory interfaces. High spatial individuals outperformed low spatial individuals only when information search tasks required the use of spatial ability in mentally constructing a model of the organisation and structure of embedded task information. (Stanney and Salvendy, 1995). Individuals with low spatial ability are more directed to the semantic content. If the semantic content of the interface is low, individuals of high spatial ability have a performance advantage compared to individuals of low spatial ability. In these situations individuals of low spatial ability derive comparatively greater benefit from the provision of additional non-spatial semantic information (Westerman, 1995).

Summary

There are some developmental evidence for a multidimensional view of spatial cognition, with different ages-of-onset for different aspects of the spatial ability. Differences in the spatial ability could also, as a complementary approach, be explained in terms of socio-cultural differences. The largest sex differences in favour of males can be found on the mental rotation aspect of spatial cognition. Recent research indicates that these sex differences could, however, be a consequence of socio- cultural differences, as for instance gender roles. They are, furthermore, of a small order of magnitude, accounting for only 1% to 5% of the variance.

Route memory may involve both scene and layout representation, which may be differentially sensitive to age. Age differences are found in configural and integrative aspects of spatial memory and could be explained in terms of older adults disposition towards a more personal and less holistic view on the environment, and/or due to limitations in working memory capacity. Older adults may benefit more from explicit verbal instructions to focus attention on the path or route than from being provided with a model or an analogy over the space. Different strategies of using landmarks and turns could promote the acquisition of other aspects of spatial knowledge. Learning a new environment is constrained by stage-based processes where route knowledge may be dependent upon sufficient landmark knowledge, and configural knowledge may be dependent upon sufficient route knowledge.

Improvements in performing spatial activities are task specific. It is difficult to study for how long these effects lasts, because there is a huge amount of variables that are difficult to control for, and which could effect the subjects performance. The acquisition of a richer knowledge base and the improvement of general reasoning ability will also effect the spatial ability, therefore it is difficult (if the task specific effect lasts) to know if it is really the task specific practise that is lasting or if it is an over all improvement in the general reasoning ability. Intellectual abilities are specific to a particular domain or function whereas a cognitive style cuts across domains. The cognitive style field-independence Vs field-dependence have effect on map retention. Field-independent learners remember more of a map's features than learners who are field-dependent. In the use of computers, field-dependent subjects are less likely to know the command, and more likely than field-independent subjects to ask for help without making any attempt at the task. The personality dimension introversion/extraversion may be suitable for determining whether an individual will be drawn toward a computer-related career. Individuals who are intuitive and thinking exhibit lower computer anxiety. These individuals are familiar with theoretical and abstract issues and attempt to make generalisations from a logical base

ACKNOWLEDGEMENTS

Thanks to all participants in the PERSONA group for feedback and to HSFR (The Swedish Council for Research in the Humanities and Social Sciences) which also founded this work.

REFERENCES

Baddeley, A. D. (1986). *Working memory.* Oxford, England, Clarendon Press.

Baenninger, M., & Newcombe, N. (1989). "The Role of Experience in Spatial Test Performance: A Meta-Analysis." *Sex Roles Vol. 20(Nos. 5/6).*

Bostic, J. Q., & Tallent-Runnels, M. K. (1991). "Cognitive Styles: A Factor Analysis of Six Dimensions with Implications for Consolidation." *Perceptual and Motor Skills Vol. 72:* 1299-1306.

Caplan, L., J. & Schooler, C. (1990). "The Effects of Analogical Training Models and Age on Problem-Solving in a New Domain." *Experimental Aging Research Vol. 16(No. 3):* 151-154.

Chu, P. C., & Spires, E. E. (1991). "Validating the Computer Anxiety Rating Scale: Effects of Cognitive Style and Computer Courses on Computer Anxiety." *Computers in Human Behavior Vol. 7:* 7-21.

Coventry, L. (1989). "Some effects of cognitive style on learning UNIX." *Int. J. Man-Machine Studies Vol. 31:* 349-365.

Craik, F. I. M., & Lockhart, R. S. (1972). "Levels of processing: A framework for memory research." *Journal of Verbal learning and Verbal Behavior Vol. 11:* 671-684.

Duesbury, R. T., & O'Neil, Jr., H. F. (1996). "Effect of Type of Practice in a Computer-Aided Design Environment in Visualizing Three-Dimensional Ojects From Two-Dimensional Orthographic Projections." *Journal of Applied Psychology Vol. 81(No. 3):* 249-260.

Furnham, A. (1992). "Personality and Learning Style: A Study of Three Instruments." *Person. individ. Diff. Vol.13(No. 4):* 429-438.

Gauvain, M. (1993). "The Development of Spatial Thinking in Everyday Activity." *Developmental Review Vol. 13:* 92-121.

Gittler, G., & Vitouch, O. (1994). "Empirical contribution to the question of sex-dependent inheritance of spatial ability." *Perceptual and Motor Skills Vol. 78:* 407-417.

Greenfield, P. M., Brannon, C., & Lohr, D. (1994). "Two-Dimensional Representation of Movement Through Three-Dimensional Space: The Role of Video Game Expertise." *Journal of Applied Developmental Psychology Vol. 15:* 87-103.

Hamilton, C., J. (1995). "Beyond sex differences in visuo-spatial processing: The impact of gender trait possession." *British Journal of Psychology Vol. 86:* 1-20.

Hasher, L., & Zacks, R. T. (1979). "Automatic and effortful processes in memory." *Journal of Experimental Psychology: General Vol. 108:* 356-388.

Hyde, J.S. (1981). How large are cognitive gender differences? *American Psychologist*, 36, 892-901.

Kolb, B., & Whishaw, I. Q. (1990). *Fundamentals of Human Neuropsychology.* New York, W. H. Freeman and Company.

Lawton, C., A. (1994). "Gender Differences in Way-Finding Strategies: Relationship to Spatial Ability and Spatial Anxiety." *Sex Roles Vol. 30(Nos. 11/12).*

Linn, M. C., & Petersen, A. C. (1985). "Emergence and characterisation of gender differences in spatial abilities: A meta-analysis." *Child Development Vol. 56:* 1479-1498.

Lipman, P. D. (1991). "Age and exposure differences in the acquisition of route information." *Psychology and Aging Vol. 6:* 128-133.

Lipman, P. D., & Caplan, L. J. (1992). "Adult Age Differences in Memory for Routes: Effects of Instructions and Spatial Diagram." *Psychology and Aging Vol. 7(No. 3):* 435-442.

Lohman, D. F., & Nichols, P. D. (1990). "Training spatial abilities: Effects of practice on rotation and synthesis tasks." *Learning and Individual Differences Vol. 2(No. 1):* 67-93.

Magliano, J. P., Cohen, R., Allen, G. L., & Rodrigue, J. R. (1995). "The impact of a wayfinder´s goal on learning a new environment: Different types of spatial knowledge as goals." *Journal of Environmental Psychology Vol. 15:* 65-75.

McClurg, P. A., & Chaillé, C. (1987). "Computer games: Environments for developing spatial cognition?" *J. Educational Computing Research Vol. 3(No. 1).*

Messick, S. (1994). "The Matter of Style: Manifestations of Personality in Cognition, Learning, and Teaching." *Educational Psychologist Vol. 29(No. 3):* 121-136.

Morris, D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior Vol. 16:* 519-533.

Pezdek, K. (1983). "Memory for Items and Their Spatial Locations by Young and Elderly Adults." *Developmental Psychology Vol.19(No. 6):* 895-900.

Pocius, K. E. (1991). "Personality Factors in Human-Computer Interaction: A Review of the Literature." *Computers in Human Behavior Vol. 7:* 103-135.

Puglisi, J. T., Park, D. C., Smith, A. D, & Hill, G. W. (1985). "Memory for two types of spatial location: Effects of instructions, age, and format." *American Journal of psychology Vol.98(No. 1):* 101-118.

Rolfhus, E. L., & Ackerman, P. L. (1996). "Self-Report Knowledge: At the Crossroads of Ability, Interest, and Personality." *Journal of Educational Psychology Vol. 88(No. 1):* 174-188.

Rosser, R. (1994). "The Developmental Course of Spatial Cognition: Evidence For Domain Multidimensionality." *Child Study Journal Vol.24(No. 4).*

Salthouse, T. A., Mitchell, D. R. D., & Palmon, R. (1989). "Memory and Age Differences in Spatial Manipulation Ability." *Psychology and Aging Vol. 4(No. 4):* 480-486.

Schiano, D. J., Cooper, L. A., Glaser, R., & Zhang, H. C. (1989). "Highs Are to Lows as Experts Are to Novices: Individual Differences in the Representation and Solution of Standardized Figural Analogies." *Human Performance Vol. 2(No. 4):* 225-248.

Schwartz, N. H., & Phillippe, A. E. (1991). "Individual Differences in the Retention of Maps." *Contemporary Educational Psychology Vol. 16:* 171-182.

Shah, P., & Miyake, A. (1996). "The Separability of Working Memory Resources for Spatial Thinking and Language Processing: An Individual Differences Approach." *Journal of Experimental Psychology: General Vol. 125(No. 1):* 4-27.

Sharps, M., J., Price, J., L. & Williams, J. K. (1994). "Spatial Cognition and Gender, Instructional and Stimulus Influences on Mental Image Rotation Performance." *Psychology of Women Quarterly Vol. 18:* 413-425.

Stanney, K. M., & Salvendy, G. (1995). "Information visualization; assisting low spatial individuals with information access tasks through the use of visual mediators." *Ergonomics Vol. 38(No. 6):* 1184-1198.

Strong, G. W., & O'Neill Strong, K. E. (1991). "Visual guidance for information navigation: a computer-human interface design principle derived from cognitive neuroscience." *Interacting with Computers Vol. 3(No. 2):* 217-231.

Stumpf, H., & Eliot, J. (1995). "Gender-related differences in spatial ability and the k factor of general spatial ability in a population of academically talented students." *Person. individ. Diff. Vol. 19:* 33-45.

Tiedemann, J. (1989). "Measures of Cognitive Styles: A Critical Review." *Educational Psychologist Vol. 24(No. 3):* 261-275.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). "Magnitude of Sex Differences in Spatial abilities: A Meta-Analysis and Consideration of Critical Variables." *Psychological Bulletin Vol. 117(No. 2):* 250-270.

Weller, H. G. (1995). "Improving the Effectiveness of Learning Through Hypermedia-Based Instruction: The Importance of Learner Characteristics." *Computers in Human Behavior Vol. 11(No. 3-4):* 451-465.

Westerman, S. J. (1995). "Computerized information retrieval: Individual Differences in the use of spatial Vs nonspatial navigational information." *Perceptual and Motor Skills Vol. 81:* 771-786.

Chapter 6
Social Navigation

Martin Svensson
SICS

A common way to navigate an information space in the real world is with help of other people, e.g. instead of looking at my map I ask another person for the direction of the place I want to go to. This communication with other agents (human or artificial) to navigate an information space is what we call social navigation.

In this paper social navigation is examined; how does it work; who uses it; and most importantly, how do we navigate socially in virtual information spaces? Basically it is possible to identify two types of social navigation: direct social navigation, and indirect social navigation.

Social Navigation

Martin Svensson
SICS Box 1263 S-164 Kista Sweden
martins@sics.se
http://www.sics.se/~martins

A common way to navigate an information space in the real world is with help of other people, e.g. instead of looking at my map I ask another person for the direction of the place I want to go to. This communication with other agents (human or artificial) to navigate an information space is what we call social navigation.

In this paper social navigation is examined; how does it work; who uses it; and most importantly, how do we navigate socially in virtual information spaces? Basically it is possible to identify two types of social navigation: *direct social navigation*, and *indirect social navigation*.

Introduction

In the new information society people are flooded with information, and to navigate these huge amounts of information is a daunting task. For some people navigating in information spaces is a relatively straightforward task. However for various reasons the same task to other people may be substantially more difficult. It is important to find ways of aiding people that have problems navigating within information spaces.

In this paper I will primarily address issues relating to social navigation. In order to understand social navigation a short summary of two related concepts: *information spaces* and *navigation* will also be given. Information spaces will be discussed later in this paper and for now it is sufficient to know that an information space is anything that allows information to be stored, received, and possibly transformed (Benyon and Höök, 1997). Information spaces can be physical or virtual; the physical information spaces build up the world we inhabit.

Navigation is the activity of 'locating' places in an information space. Navigation can be separated into two distinctive activities: exploration, and wayfinding. In wayfinding the navigator has a specific destination they want to travel to, and in exploration the navigator simply explores the information space more or less randomly. Downs and Stea (1973) defines wayfinding as an activity composed of four steps:

1. Orienting oneself in the environment
2. Choosing the correct route
3. Monitoring the route
4. Recognising that the destination has been reached

Moving around in an information space does not automatically mean that we navigate within the space. For example, the first time someone goes to work they are navigating, however after a while they know the route without having to think and eventually they can transport themselves to work, navigation is a cognitively demanding activity. In the same way there is a difference between being lost in a space and to explore a space. When a user is lost she has no sense of location, and no ability to monitor a given route, and by introducing social navigation we try to attack the problem of "being lost in space".

Several studies have shown that certain users have problems when navigating in different information spaces (Höök, 1996). Individual differences such as spatial ability and technical aptitude are key factors in determining if a user will be successful in navigating within an information space. One way to deal with these differences is to incorporate various metaphors to structure the information in the information space (Dieberger, 1994). For example, a library is known to most people and could therefore serve as an excellent metaphor when structuring certain kinds of information spaces. Yet another example is Dieberger's (1994) virtual city, where information is structured around concepts such as streets, houses, landmarks, etc. These spatial metaphors make it easier for users to build cognitive maps over the information space, and thus reduces the risk of getting lost. However, spatial metaphors are not enough, as some users will keep getting lost. In a city and hence a virtual city, people may get lost, and others will never find a book in a library without help from a librarian. These people use other means of navigation to find their way through an information space. People that have a hard time reading maps usually end up asking their way around a city or following crowds of people. For example in an airport, where there are several navigational tools such as signs, maps, bulletin boards, etc., a common way to find the baggage claim point is to follow the "crowd", ignoring the different signs or maps that tell how to get to it.

This direct and indirect interaction with other people can be thought of as social navigation, that is, in order to navigate the information space people communicate with other inhabitants of the space. In the subsequent sections I will give a definition of social navigation and in depth discuss the different ways in which social navigation can occur.

Social navigation

The social navigation metaphor (Dourish and Chalmers 1994; Erickson 1996; Dieberger 1997) is fairly new and has been used to capture certain aspects of navigation. The definitions vary from "exchanging pointer in the WWW" to "socially enhanced navigation".

The presumption used in this paper is that social navigation is navigation. In fact, following on from the above discussion on navigation, the only thing that separates navigation and social navigation are the tools. Further to this, to understand social navigation we need to extend our navigational tools to incorporate social elements. Therefore when we navigate socially there is one more tool available to us, namely an agent or a group of agents that can help us navigate within an information space. In addition other non-social tools are still available to a user, e.g. if a person is lost in a city they can ask another person (the social tool) to point out their destination on a map. Having done so they can then use the map (the non-social tool) to find their destination. Similarly, in a virtual information space a user can ask an agent for a specific destination and then use landmarks to navigate to it. Perhaps it is more appropriate to think of social navigation as socially enhanced navigation. For clarity I will divide social navigation into two categories:

- Direct social navigation
- Indirect social navigation

The agent has several characteristics and properties. Firstly, the agent can either be artificial or human. Secondly, a user can either communicate directly or indirectly with the agent to navigate in the information space. A user communicates indirectly when there is no mutual communication between the agent and the user. In a museum, for example, a guided tour would be classified as indirect social navigation if the guide (agent) takes the spectator (user) around the museum, and there is nothing the spectator can do to alter the route or pace of the tour. However, there is no clear cut line between indirect and direct social navigation, for instance, the guided museum tour can also be thought of as direct social navigation if the spectator (user) has the ability to influence the tour. Thirdly, depending on what type of information space and navigational task, the agent(s) can be of help in different ways. In a complex information space such as a word processor, a user often has vague ideas on where they want to go. Further to this if they had the ability to communicate directly with the agent, the agent could pose questions and try to clarify the user's goal or destination.

In the following discussion these properties and characteristics will be discussed in greater detail. However lets start off by defining the different types of information spaces that social navigation can be applied to.

Information spaces

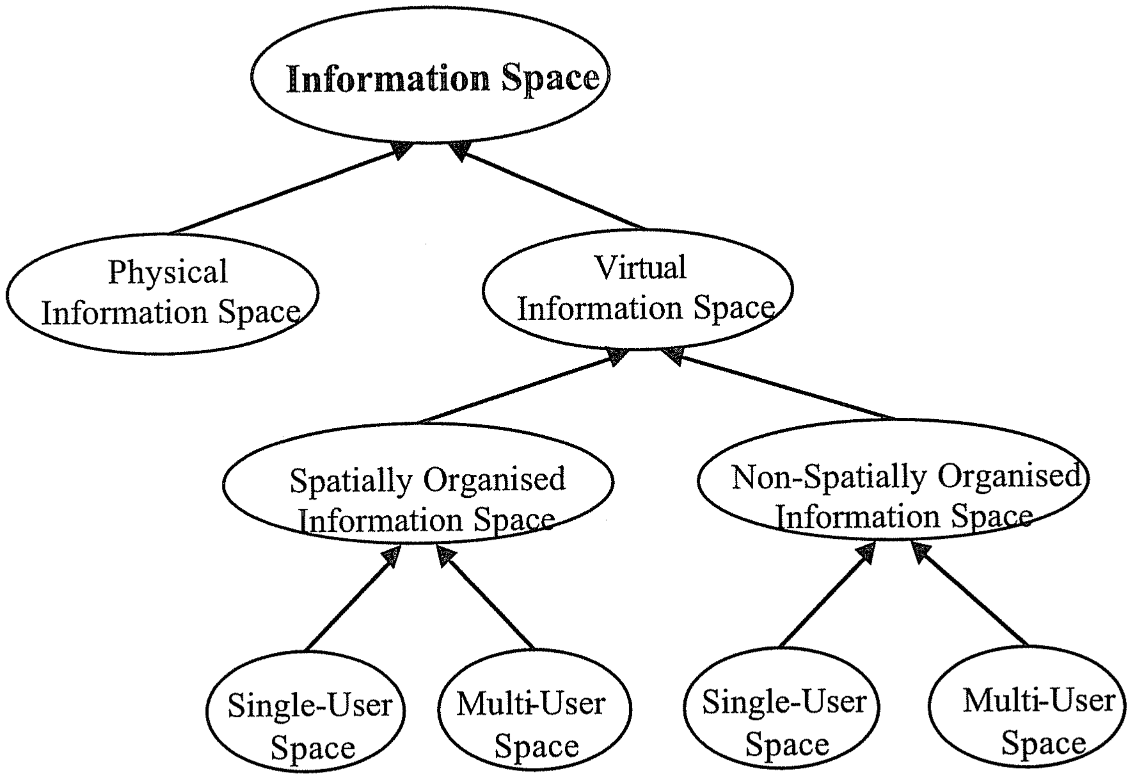


Figure 1. Information Spaces

An information space is anything that allows information to be stored, retrieved, and possibly transformed. They are either physical or virtual. Physical information spaces build up the world that we inhabit, and virtual information spaces are those typically found in computers and networks of computers. For example, the file structure on the local harddrive and the WWW (World Wide Web) are examples of virtual information spaces. It is possible to find information spaces within other information spaces, e.g. on my harddrive I have a word-processor, which constitutes an information space with specific properties. In this paper the focus is on virtual information spaces, how we navigate these, and how social navigation can enhance navigation in these spaces.

A virtual information space can either be spatially or non-spatially organised. Furthermore, information spaces can be single-user spaces (a word processor) or multi-user spaces (WWW).

In spatially organised information spaces (Dourish and Chalmers, 1994) it makes sense to talk about properties such as up/down, north/south, left/right, etc. Physical information spaces are of this kind and in a sense virtual information spaces that are spatially organised tries to mimic the real world, or certain aspects of it. Dieberger's virtual city, for example, is an attempt to organise an information space around concepts such houses, streets, etc., thus, allowing a user to navigate the information space much in the same way that she would navigate a city. Another metaphor taken from the real world that can be found in almost every PC is the desktop metaphor. Here the information space is organised as a desk, with properties such as a trashcan, folders, etc.

There are basically two categories of environments that use spatial models for multi-user information spaces:

- MUD environments
- Virtual Reality environments

MUDs (Multi User Dungeon) are multiple-user network games. A MUD is a text-based virtual world, whilst virtual reality environments are 3D based virtual worlds.

In comparison, non-spatially organised information spaces are not built up by concepts from the physical world. Information in these spaces is semantically organised (Dourish and Chalmers, 1994) where related information is grouped together. Hypertext systems are built up this way, with links between related information that build up the hypertext system. In non-spatially organised information systems (such as the WWW) many of our ordinary navigational tools and skills is of no use to us. It is hard to talk about north and south, in effect, making compasses obsolete, landmarks are also difficult to incorporate in these spaces, and finally the physical distance between nodes seems irrelevant, i.e. the common way to measure distance is to count the number of nodes visited. These spaces can both be single-user and multi-user spaces. The WWW is a typical multi-user hypertext system, and the online help for a word-processor is a typical single-user hypertext system.

To summarise four types of virtual information spaces are of interest to us:

- spatially organised single-user information space
- spatially organised multi-user information space
- non-spatially organised single-user information space
- non-spatially organised multi-user information space

In the spatially organised information spaces it is easier to use navigation skills and tools taken from the physical world, especially those that are found in multi-user spaces. Conversely information space such as the WWW navigation is a complicated task, the reason being that the WWW is not homogenous, i.e. objects not related to each other are linked. Maps are virtually non-existent and the chance of a user not finding their destination is high.

Lastly spaces can either be open (dynamic) or close (static). Closed spaces are characterised by the fact that they are stable in time, i.e. they seldom expand or collapse. On the other hand open spaces can substantially change over time, for example, the WWW is an ever growing space. Open spaces are harder to navigate than closed spaces, and the tools that can be used to support navigation in those spaces are fewer than in the closed ones. For instance, suppose that we want to use an agent that explains to a user how the information space is organised: relations between nodes, where to find certain information, etc. In a dynamic space the agent can quickly turn obsolete and instead of helping the user navigate the information space it will probably hinder them. Search engines on the WWW have this problem, they often give users information about nodes that no longer exist, and this is of course due to the dynamic structure of the WWW, where nodes are constantly added and removed.

Direct social navigation

Direct social navigation is characterised by the fact that there is mutual-communication between user and agent. A user can ask questions like "Where am I?" or, "Where can I find location X?" The agent answers the user and perhaps more importantly asks the user to clarify their questions. An agent can help clarify the user's goals or even change them. When users are uncertain of where they want to go the agent can support them in formulating a destination.

There are three forms of direct social navigation; firstly, a user can communicate with an artificial agent; secondly, a user communicates with a human agent; and finally, a user communicates with an agent that in turn uses other agents. Each situation has certain benefits and drawbacks. Below they are discussed in greater detail

Communication between a user and an artificial agent

This kind of social navigation to some extent already exists. Typically in MUD environments several experiments with artificial agents guiding users have been conducted. Julia (Foner, 1993) is a MUD agent or robot that a user can ask questions such as "Who are logged in the MUD?", "Where can I find location X?". Julia enjoys a good conversation; her favourite discourse is NHL hockey. In addition when she does not understand the user she will notify them about it. Her main purpose is to keep track of mud users, but she also has the ability - to some extent - formulate a user's goals. Julia has been shown to be very successful in helping users that are lost in the MUD.

Two examples of conversations with Julia (Foner, 1993):

TheHighMage says, "Julia, I'm bored. Where should I go?"
Julia says, "You should go see gilded door, TheHighMage."
Julia says, "From here, go present, out, north, n, w, n, up, up, 7, up, up, s."

You whisper, "where are we?" to Julia
Julia whispers, "We are in "The Inter Nexus""

Some interesting properties arise in MUDs or in virtual information spaces in general. For instance, consider this following example (again taken from Foner):

You say, "julia, where is leira"
Julia says, "Leira was in the Inter Temporal Chat Lounge about 6 weeks ago."

The above example demonstrates one of the major advantages with artificial agents. An artificial agent is extremely good at storing information and retrieving it quickly. A human agent is not capable of doing this. In extremely dynamic information spaces an artificial agent could be of great help; it is relatively easy for an artificial agent to keep track of history, and can thus in a graceful way inform a user that some piece can no longer be found in the information space. The WWW is a prime example of the problems which can arise in dynamic environments because pages are frequently moved or deleted. The Alexa system (http://www.alxa.com) keeps an internal history list with previously visited nodes, so when a user comes across a non-existent node Alexa retrieves that page from the internal database. However, the storing of these non-existent semantic relations between web pages may not help the user in navigating the WWW.

The MS Word97 agent (http://www.microsoft.com) is an attempt to implement this kind of social navigation in a single user environment. The agent allows the user to type questions in natural language and it monitors the user's actions so it can give the right answer depending on the user's task at hand. However, the MS Word agent fails to implement social navigation in one major way, there is no mutual communication between the user and the agent. This may appear to be a contradiction after stating that the user is allowed to pose questions in natural language however the problem lies elsewhere. Typically agents are not capable of interacting on the same level as the user. In my opinion, allowing a user to interact with the agent in natural language creates uncalled for expectations when the agent is not capable of doing the same. In fact, to my knowledge no successful agent like Julia has been implemented in a single-user environment such as a word processor, or window manager. To be able to support social navigation in this way it is necessary to make certain restrictions to the agent.

Firstly, the agent has to work in a very restricted domain or information space. Take Julia for example, she works in a closed information space with all information about the MUD accessible to her at any given time – the only thing that changes are the users logging in and out.

Secondly, it is necessary that the agent has some knowledge of the user. How experienced is the user with the information space. For example what are their preferences? What goals are they trying to achieve? To some extent the problem with the MS Word is its inability to distinguish an experienced user from a novice. The success of MUD agents can be traced back to the fact that there is a very limited number of tasks that users can undertake, i.e. no matter how experienced a user is in their tasks, they will not be that different from a novice user.

Lastly, it is important that the user has confidence in the agent. A user's trust in the agent is of course a measurement on the performance of the agent. If the agent always leads the user onto the wrong track, the trust in the agent will decrease. I believe that user's trust in an artificial agent is generally lower than in a human agent. We are in general more forgiving if a human makes a mistake than if a program makes one. This fact makes it extremely important that the artificial agent gives good advice, rather than giving lots of advice. In the information retrieval society this is known as precision and recall. Precision is a measurement of the relevance of the advice. Where as recall measures the number of relevant advises given, compared to the total space of relevant advises. Amant and Dulberg (1998) showed that agents who help users to navigate should aim for precision rather than recall. However, they also found out that having an agent that gives advise (although poor) is better than having no agent at all. They explain this fact as being due to the users ability to discriminate good from poor advice.

Communication between a user and a human agent

Let us now turn to social navigation that involves human agents. This type of social navigation is suitable for multi-user environments such as the WWW or MUDs. It can, however, also occur in single-user environments that are, in one way or the other, connected to multi-user environments. For example, we can imagine a scenario where a user is stuck in their spreadsheet. The user, instead of consulting the spreadsheet's online manual, can send out a request for help to a pool of spreadsheet experts on the WWW. An expert answers the user's request and a real time communication begins to navigate the user to their goal. In this scenario it is also likely that the expert will help the user reformulate their goal, as it is often the case that we have vague ideas of what we are trying to accomplish.

As with artificial agents it is important that a user trusts the human agent. When I (as a user) consult a human agent, I need some sort of guarantee that the agent is the expert that they claim to be. MUDs solve this problem by introducing the concept of players and wizards. Wizards are expert MUD players and can only be granted their wizard status by other wizards. Assuming that wizards do their job properly, a user can be sure of getting accurate answers from them.

There are basically two ways of identifying experts. A system can either test the expert's knowledge in an automated fashion, e.g. the SATELIT system (Akoulchina and Ganascia, 1997) or other experts can grant expert status, e.g. becoming a wizard in a MUD or an official guide in the PowWow system. Which ever is preferable depends on the domain. In extremely complex and large domains it may be difficult to find good automated ways of identifying true experts. However, whenever it is possible an automated method should be preferred.

In the PowWow system (http://www.tribal.com) there are a number of official guides that help novices. The guides can be found in certain communities[1]. In the following transcript from a PowWow session, the user "swede" could not figure out how to view a personal profile. The PowWow online help did not give him much assistance, so he consulted one of the online guides:

laila>	*I see you don't have a profile swed*
	"swede use the online help to figure out how to view a profile"
swede	*>laila, how do I see another persons profile*
laila>	*Ask GURADIAN he is here to help*
swede>	*ok*
swede>	*GUARDIAN how do I get another persons profile*
GUARDIAN>	*swede, click on the community button, in your personal communicator*
swede>	*ok*
GURADIAN>	*now right click on the person whose profile you want to see*
swede>	*GUARDIAN, nothing happens*
GUARDIAN>	*ok....what version of pw do you have...you can find version number under "About PowWow..." in the "help" menu in your personal communicator*
swede>	*3.1, GUARDIAN*
GUARDIAN>	*ok I see...you nee version 3.2...go to* http://www.tribal.com/download *to get it*
swede>	*thanks!*

Another example of this direct communication between a user and a human agent is to navigate the information space together. In the following example the users freddie and swede are engaged in a conversation. User freddie asks swede if he knows any interesting web-sites:

freddie>	*swede to you know any cool sites*
swede >	*just hold on and I'll show you*
	"swede sends a request to freddie to join in a web cruise"
	"freddie accepts"
	"swede loads a HTML-page in his browser"
swede>	*is anything happening with your browser*
freddie>	*yes it loaded a HTML-page*
swede>	*good, now let me show you around...*
	"Now swede acts as freddie's human agent"

The question is what makes it possible to classify the above examples as social navigation? In the first example the user "swede" is clearly trying to move from one location to another. He wants to move from his current location, to the location from where he can see the personal profile of a user. In order to navigate the information space (in this case the PowWow personal communicator), his navigational tool is a human agent. He can at any time consult the agent to get help on where he is in the information space. He also uses the agent to find his destination and perhaps more interestingly to change his goals. His intention was at first to view another person's personal profile, but he ended up installing a new version of PowWow. His new goal is of course only a means to the higher end of viewing a person's profile.

[1] A community is a chat room that discusses a certain topic. For example, the PowWow community "New PowWow Users"

In the second example there is no clear destination. Here the user "freddie" uses his agent as a means to not getting lost in the information space, where swede can be thought of as the navigator on a ship. The user "freddie" explores the information space (WWW) with help of his agent, the hope is that the agent will take him to some interesting places. We can think of this situation as exploring the archipelago with a compass and chart, without having any predefined destination.

As mentioned earlier, artificial agents are very good at keeping a history record of what is happening in the information space. They are, however poor in helping a user on things outside their own information space, this is in contrast to human agents which are much better at this. In the above example with "swede" and "GUARDIAN" an interesting phenomenon arise. The human agent ("GUARDIAN") knew that another information space (PowWow 3.2) had the information that "swede" wanted, which the online help did not know. This may not appear to be a big problem, but with more and more people doing collaborative work, exchanging documents, etc. different versions of the same software (information space) can cause huge compatibility problems. This brings us to the final type of direct social navigation which occurs, namely communication between a user and several agents.

Communication between a user and several agents

Navigating socially with help of several agents is a combination of the two examples described above, and what I believe to be the most fruitful and interesting form of direct social navigation. There are both human agents and artificial agents that aid the user in navigating within an information space. In the user's view there is only communication with one agent, either human or artificial, which serves as an interface agent between the user and the other agents. This is not to say that a user cannot utilise several interface agents, what it means is simply that a user only communicates with one agent to solve a specific task, although the agent in turn uses other agents. Using several agents (both human and artificial) has a number of advantages to using a single agent. Firstly, it is easier to connect a single-user environment to a multi-user environment, where an interface agent provides the link. This compares favourably to making the user themselves having to connect to the multi-user environment, for example "swede" and "GUARDIAN" above. An interface agent in the single-user environment could do this, transparent to the user. For example, when a user searches the online help in a spreadsheet, there can be an option to consult other agents when needed. Secondly, more information spaces are accessible to the interface agent. It is possible to imagine a word processor agent that has access to a pool of agents (earlier and later versions), in effect, it can give answers such as "This can be found in a later version of your word processor". Thirdly, it is easier to design a system where an agent communicates with another one, where the latter contains comprehensive domain specific knowledge. The approach suggested would appear to be superior to having the situation where one agent carries out all the tasks. That is, it is desirable to create small but specialised agents that collaborate. This increases flexibility and reusability of agents. Lastly, utilising small and specialised agents has a positive effect on trust. Instead of having one general agent that tries to answer every question, a specialised agent can inform the user that the question is outside the agent's domain, and possibly redirect the user to another agent.

Indirect social navigation

As stated earlier in indirect social navigation there is no two-way communication between the user and agent. Some of the techniques presented below may seem a little far-fetched to call it social navigation as we have defined it, i.e. a user communicates with an agent or agents to navigate an information space. However, it is important to include these techniques as they use different types of social phenomena to aid the user in their navigational task

Filtering agents

There is information all around us, in newspapers, on television and radio, etc., the information is supplied to us on a constant 24 hour basis. How do we manage to process all this information? We use various techniques to filter the information; we let editors filter the information in newspapers; we listen to the radio stations that we like; we allow friends with similar tastes as ours to recommend movies; and finally, we learn by trial and error where to find good information. With the introduction of the Internet and the WWW many of our ordinary tools to filter out irrelevant information were pulled away from us. A common technique to find good information on the WWW is to use different search engines and hopefully come across relevant information, a tiresome and time-consuming way to find information. Several methods of automatically separating good information from bad information based on personal preferences have evolved. Two such techniques are content based filtering (Malone et al 1987) and collaborative filtering or social filtering (Miller, Riedl, and Konstan, 1997; Shardanand and Maes, 1995).

In content based filtering an agent scans the information for specific parts of it that match some criteria, and based on some statistical analysis rates the relevance of the information to the user. Usually keyword-matching techniques are used to filter the information, that is, the user supplies a preference file to the agent with keywords that the agent shall look for in documents. For example, if I am interested in social navigation I could have an agent that on regular basis scans a newsgroup for documents that contain the words navigation, collaborative, and social. Shardanand and Maes identifies three major drawbacks with these kind of filtering methods:

- Items must be parsable, e.g. music files can not be parsed using content based filtering
- It is difficult to achieve serendipity
- Content based filtering disregards aspects such as quality or style

In addition the success of the filtering is highly dependent on the keywords the user supplies to the filtering agent. In the above example my filtering agent would probably recommend many documents that have nothing to do with social navigation, this is due to the keywords I have picked; navigation, collaborative, and social are words that can be found in many discourses.

To overcome the problems with content based filtering, social filtering has been proposed. Social filtering recommends information based on what other people with similar tastes like or dislike. Social filtering is relatively straightforward to implement; all users of the system are connected to a server that keeps track of every user, users can be anonymous and personal profiles of each user are stored. Personal profiles are matched and the system creates clusters of users with similar tastes. Therefore whenever a user comes upon a new piece of information they can see what other people with similar interests as them think of that particular piece of information. If a social filtering system is to work some sort of rating of the information pieces has to be done by the users of the system. This is so the system can create and cluster personal profiles. The more people that utilise rating system, the more accurately the system can group users. Ratings can be done either explicitly and/or implicitly; implicit ratings are, for example, time spent reading an article, while on the other hand explicit rating let users score information pieces.

The major benefits with social filtering are that it enables the agent to filter information based on quality. The filtering is not based on content, but on other user's ratings of it. Furthermore, it allows for recommendations of information pieces that the user has not asked for earlier. For example, a social filtering system can tell me that some information (that I have never asked for) may be of interest to me; the agent draws this conclusion based on what other people, with similar interests as me, have said about it.

There are two major problems with social filtering systems. Firstly, social filtering systems suffer from *cold starts*. Therefore, in order for a social filtering system to work correctly it needs some sort of input to work with. Hence when there are no or few ratings to build the personal profiles on, the system does not make good predictions. Secondly, explicit rating of information is not all that simple. A particular problem is the snow ball effect. The snow ball effect arises when a piece of information (which is already rated highly) is rated similarly other users. As a result a snowball effect arises where a piece of information is constantly rated highly. Miller and colleagues illustrated that information with high ratings, twice as often got rated than information that was rated low. Furthermore, how do we judge ratings from a user that has created the information (e.g. in a system containing articles it is likely that I will give my own articles high ratings)?. This is a problem that is likely to occur in social filtering systems were users are anonymous.

Two systems that have been successful in using social filtering are the firefly (http://www.ffy.com) system and the GroupLens (Miller, Riedl, and Konstan, 1997) system. Both use explicit ratings to create and group personal profiles. Firefly is a WWW music recommendation system and the GroupLens system is connected to a newsreader. In the GroupLens system it was shown that social filtering indeed enhanced predictions compared to average ratings on each article.

History enriched environments

What other people have done in the past can tell us something about how to navigate the information space. If we get lost in the woods and come upon a trail, a good idea may be to follow that trail. History enriched environments use the same idea; the more people that take a certain path through the information space usually indicate that this path is a good one. History enriched environments have been used in muds to guide (novice) players in the right direction. For example, when a player enters a room in a mud the following information is given:

> *This is the entrance to the mud world "social navigation". On your right hand you can see a door that doesn't seem to have been used in ages. In front of you see a big portal that several people have stepped through in the past.*

Apparently the intention is to lead a user in a certain direction based on what other people have done in the past. Another example of history enriched environments could be to automatically change the colours on the links in a web page based on usage of links. Alexa records whenever a user of the system enters and leaves a web page. This way a user can always tell what other users of the Alexa system usually go from a web page.

One problem with history enriched environments is that they have snowball effects. Considering the Alexa system again; if only one user has visited a page before me, Alexa will recommend to me that I should follow her path – though that may very well be the wrong path to take. If I take that path, the next user will see an even bigger "trail" from that page, thus the snowball gets bigger and bigger.

Social spaces – World Wide Web

Erickson (1996) discuss the social structure of the World Wide Web. Even if there is no communication going on between an agent and a user this is important to mention. The WWW in itself can be seen as both a history enriched environment and a social filtering system. A major part of the nodes on the WWW are personal homepages, that is, pages that users of the WWW set up. They usually contain information about the person, what the person works with, how old they are and what their interests are, etc. Homepages often contain links to other (friend's) homepages and other interesting places on the WWW, creating a "social" network. Therefore if I know someone who is an expert in cinematography, a good place to start to look for relevant literature in that topic would be on that person's homepage.

Narratives

Something that has not been mentioned so far is how the agent communicates with the user, spatially, in natural language, or in some other form. The only thing that has been said is that communication occurs. I will therefore give a short summary of narratives (which I believe is a natural form of communication between people) and how narratives can be used in communication with a user; for a comprehensive discussion on narratives the reader should turn to Persson (1998).

Let us revisit the example with Julia above:

> *TheHighMage says, "Julia, I'm bored. Where should I go?"*
> *Julia says, "You should go see gilded door, TheHighMage."*
> *Julia says, "From here, go present, out, north, n, w, n, up, up, 7, up, up, s."*

The interesting part of this example lies in Julia's answer to TheHighMage's question; Julia's answer is purely spatial. The answer gives TheHighMage very specific instructions on how to go to an (in Julia's mind) interesting location but does it encourage learning the information space? Probably not, if TheHighMage in the future finds himself at the same location, he will not remember that he should "go present, out, north, n, w, n, up, up, 7, up, up, s" to find an interesting place. On the other hand, Julia could try to give the instructions as a narrative:

> *Julia says, You are in the "hall of emptiness" and I would recommend that you go to the "pleasure doom" a place where other players usually meet at 5pm. From here you have to pass a number of rooms, first turn north to the "crossroads" – it got its name due to the vast number of exists...finally head south and you will end up in the "pleasure doom"*

A narrative gives coherence to the instructions (events); thus helping the user to understand the spatial and semantic relationship between different nodes in the information space. A narrative can also let a user link events temporally, e.g. on Monday I got to work and later that day I became sick and had to go home. The user understands why the agent gives the advice. The narrative also increases the user's ability to remember the information space and to remember the specific navigational situation. For instance, TheHighMage could in the future tell his friends to watch out for the "hall of emptiness" or tell them to go to the "crossroads" if they want to explore the space.

Although narrative serves as an excellent tool for communication with a user it has to be used with care. As Persson argues, a narrative can in some situations and to some users simply be annoying. Take Julia for example, her first answer would probably be the best one to give to an advanced user who only wants to find other users.

Related work

Others have dealt with the concept of social navigation using a slightly different approach. What is typical of their view of social navigation is that it takes place in virtual multi-user information spaces. For example, in hypermedia systems such as the WWW, different individuals create the nodes. They also stress the non-mutual type of communication as opposed to the more mutual communication that characterise social navigation in the real world. Below a brief description of Dieberger's, Dourish & Chalmer's, and Erickson's view on social navigation is given.

In Dieberger's view (1997) social navigation is a navigational behaviour on the Internet that is based on communication and interaction between users. Especially the phenomenon of sharing pointers in WWW can be characterised as social navigation. Sharing of pointers can be done in a variety of ways, e.g.:

- sharing of hotlists with friends and colleagues
- web pointer lists that users keep in their homepages
- emails

This type of non-mutual interaction between users characterises social navigation, that is, the information provider does not have to be aware of the information seeker. However, it is of course the case that the communication can be mutual. In Dieberger's Juggler system a MUD is connected to a browser, and when a user types in a URL the system recognises it and loads it in the browser. Another example is when a user knows that a person is an expert in a certain area; she can ask that person for interesting links in that topic. Dieberger also includes history-enriched environments as supporting social navigation.

In Dieberger's view true social navigation exists when there is URL recognition for both the sending and receiving party. That is, in the Juggler system the sender has to type in the URL and send it. The Juggler client at the other end recognises the URL and loads it into the browser. When the sender only has to click a link in their browser and the URL is transmitted to the receiver, then true social navigation arises (cf. the example with "swede" and "freddi" above).

Dourich & Chalmers (1994) define navigation as "the means by which a user can describe movement between pieces of information", and when information systems support collaborative activity social navigation arises. There are two forms of social navigation; one takes place in spatially organised information systems, and the other takes place in non-spatially organised information systems. In spatial information spaces the activity of another user or group of users can provoke social navigation. For example, moving towards people or selecting objects because someone else has examined them. The second type is characterised by information systems such as the WWW, where "navigation is not based on either location or content, but rather on recommendation and other social factors".

Collaborative filtering can also be classified as social navigation. Users can, for instance, rank information. A user can then issue queries such as; give me the top most popular articles in subject X.

Erickson (1996) views the WWW as a social hypertext. WWW is a hypermedia system that is partly built up by homepages that are used to describe people that use WWW. The nodes in WWW represent people and when users navigate from one node to the other, it can be thought of as a form of social navigation.

Erickson argues that this type of navigation strategy has a number of interesting features. For instance, users can (instead of querying a search engine) pose the question: "who would know?". Another property that is significant for social navigation is the non-mutual communication between the information seeker and information provider. As Erickson puts it "an important difference is that on the Web I can find out about what people are doing and writing without becoming obligated to them". What Erickson means is that if a person is requesting information from another person, that person is in a sense obligated to thank the other person. On the other hand, if a user visits a homepage she can use the information on the homepage, without becoming in debt to the creator of that homepage.

Concluding remarks

Concluding the discussion there are two prime components that we have to look at in order to decide whether an activity can be classified as social navigation or not. In my opinion these components are first of all the activity in itself, and secondly, the tools we use to accomplish this activity. First and foremost social navigation is navigation, therefore anything that classifies as social navigation should also be classifiable as navigation. This relationship does not hold in the other direction, if that were the case social navigation would be nothing but a synonym for navigation. Clearly it is something else that makes social navigation unique, and in my opinion it is the tools we use to navigate an information space.

A great deal of effort has been devoted to discuss and classify these tools. The tools (or some of the tools) at our disposal in social navigation incorporate some sort of social behaviour. In the paper this has been referred to as communicating directly or indirectly with an agent that can be either human or artificial. On the one hand the direct form of social navigation can best be described as a user and agent engaging in a "discussion" to solve a navigational task. While with indirect social navigation a user and agent do not have the same opportunities to interact with each other, e.g. a user can follow traces of other users; or perhaps be taken on a guided tour without having the ability to influence the guide (agent).

On top of this direct and indirect communication with agents we have discussed narratives as a form of communication technique with the user. Instead of giving a user spatial answers such as "right, right, left, left". The agent should, instead, give a user answers that make sense and that they can relate to. For example, "on your right hand you see a big church with two towers - the only one in Stockholm – anyway, there you turn left and follow the street to a park, at the end of the park turn left again, etc., etc.".

I believe that social navigation serves as an excellent metaphor to people who have difficulties in navigating spatial or semantic information spaces. Social navigation captures certain tools/aspects of navigation that we use in our every day life, and should be exploited further as an additional method to support users in navigating the huge and complex information spaces that surrounds them.

References

Akoulchina Irina and Ganascia Jean-Gabriel (1997). SATELIT-Agent: An Adaptive Interface Based on Learning Interface Agents Technology, *Proceedings of the Sixth International Conference on User Modelling (UM'97)*, Chia-Laguna, Sardinia, Italy, Juin, pp.21-32.

Robert St. Amant and Martin S. Dulberg (1998). An Experiment with Navigation and Intelligent Assistance. *Proceedings of the International Conference on Intelligent User Interfaces (IUI'98)*, pp. 171-178.

Benyon, David, and Höök, Kristina (1997). Navigation in Information Spaces: supporting the individual, In S. Howard, J. Hammond, and G. Lindegaard (eds.), *Human-Computer Interaction, INTERACT'97*, Chapman & Hall.

Dieberger, Andreas (1994). *Navigation in Textual Virtual Environments using a City Metaphor*, PhD Thesis, Vienna University of Technology, November 1994.

Dieberger, Andreas (1997). Supporting Social Navigation on the World-Wide Web, *International Journal of Human-Computer Studies*, special issue on innovative applications of the Web, Vol. 46, pp. 805-825.

Paul Dourish and Matthew Chalmers (1994), Running Out of Space: Models of Information Navigation, short paper, *HCI'94*, Glasgow, August 1994.

Downs, R. and Stea, D (1973). Cognitive Representations, pp. 79-86, in R. Downs and D. Stea (eds.) *Image and Environment*, Chicago: Aldine.

Erickson, Thomas (1996). The World Wide Web as Social Hypertext. In "Viewpoints," *Communications of the ACM*. January, 1996.

Höök, Kristina (1996), *A Glass Box Approach to Adaptive Hypermedia*, Ph.D. Thesis, Dept. of Computer and Systems Sciences, Stockholm University, ISBN: 91-7153-510-1, October 1996.

Leonard, N. Foner (1993). What's An Agent, Anyway? A Sociological Case Study, http://foner.www.media.mit.edu/people/foner/Julia/Julia.html.

Thomas W. Malone et al (1987). Intelligent Information-Sharing Systems. *CACM*, 30(5), 390-402.

Miller, B., Riedl, J., and Konstan, J (1997). Experiences with GroupLens: Making Usenet useful again. *Proceedings of the 1997 Usenix Winter Technical Conference.*

Miller, Laura (1997). *Enhancing Usability in Virtual Environments: Understanding the Cognitive Skills Required in the Use of Navigational Tools*. Ph.D. thesis, University of Central Florida.

Persson, Per (1998), Supporting Navigation in Digital Environments: A Narrative Approach. This volume.

Shardanand, U and Maes, P (1995). Social information filtering: Algorithms for automating "word of mouth". *Proceedings from CHI'95* , pages 210-217.

Chapter 7

Inhabiting Information Space: Work, Artefacts and New Realities

Alan J. Munro
Napier University

This chapter looks at three areas. It begins by discussing a number ofmainly ethnographic studies of real world workplaces, and looks inparticular at the need for people inhabiting those workplaces to be awareof what others are doing in the context of their work. A rich awareness ofwhat one's workmates are doing is shown to help those in the workplace inbeing aware of the status of the work, and neccessary for being able toform collaborative enterprises with others. It goes on to look at the kindsof artefacts used in workplaces, and discusses the use of artefacts inwork. It also discusses political issues concerning artefacts, especiallyin representing information. Finally, VR systems are studied in the lightof these discussions, in particular the representational properties ofthese systems, both in terms of representing people, and representinginformational objects. It is argued finally that in conceptualising thenavigating of invormation space we need to attend to foundational issues ofthe properties of information spaces, as well as issues of how and why wenavigate within them.

Inhabiting Information Space: Work, Artefacts and New Realities

Alan J. Munro

Dept. Computer Studies, Napier University
219, Colinton Road
Edinburgh, EH14 1JD
a.muro@dcs.napier.ac.uk

Introduction

In the Persona project, we are specifically looking at 'personal and social' navigation of information space. Thus the project is concerned with a number of interrelating things. Firstly, we are concerned with navigation per se, it could involve navigation in the real world, such as the way in which we make use of real world information in the way we use signs, the physical design of the built environment and our knowledge of our neighbourhoods to help us navigate around the world. We can take some of these concepts and apply them to other types of 'informational' navigation. A good example of this kind of navigation is that of navigating around the bookshelves in a library. In doing this, we are not just finding our way physically around shelves of books. We are actually also finding our way around a information space, around a physical embodiment of a certain classification and ordering of knowledge, so that in navigating the physical space, we are also navigating a conceptual space. We can go beyond such examples to navigation in totally conceptual spaces, 'virtual' ones, such as the various different Virtual Reality worlds and into various embodiments of information in terms of 'information terrains'.

Navigation is not the only concern of this project, however. To stop at this would give us the idea that all we are is embodied or virtual seekers in some kind of space, whether real or conceptual, and that all we have to give us information in these spaces is the various pieces of information that these spaces afford us. The other elements of this project acknowledge that this is not the case. A lot of our navigation in real spaces is done with the help of others. We may not notice this at first, but most of our navigation is accomplished with the products of others' endevours, for example in terms of street-signs, roads, paths, transport infrastructures which take us from one location to another. There are also less obvious indications to where we are and where we want to go which are provided by such things as graffiti, ad-hoc signs, etc.

Let us go further with an example about navigating in a city. In navigating around a city we may also use other types of information, which is given to us when we request help from others. A set of directions given may be diametrically different depending on our circumstances. An example of this is if a car driver and a pedestrian both ask a denizen of Oxford the way that person has to go to get to Cornmarket in the city centre from the Botley

Road. To the pedestrian, the way which will be pointed will be quite straightforward, involving walking along a fairly direct route. However, to the car driver, the directions will be radically different, giving a route which in parts heads in nearly the opposite direction from the intended destination, and which follows the best possible route for the car, through a number of one-way systems. In this example we see another important part of the persona project. Information which is given when one person helps another is often personal, and which responds to that person's situation.

We wish to take these concepts and apply them to the various types of virtual spaces which now exist, whether they be what might be argued to be an informational universe such as the World Wide Web, or whether it is a smaller scale thing, such as a database, or an interface. We have seen in the real world, that we make our way around it using others in various ways, in the virtual, informational world, we do similar things, and navigate through information, finding the sorts of things we need which might aid our various enterprises.

It has been argued that in the real world, we utilize others to aid our navigation. We do similar things in various informational worlds such as a library where we navigate in a space which is partly constructed by librarians orienting to some sort of method for arranging knowledge, where books are placed according to some sort of conceptual relations.

This chapter is intended to review some interesting work which highlights sociality and cooperative activity which can be seen to apply well to our concepts of navigation. It is taken that people will navigate in conceptual spaces for a reason, and the most likely reasons are to do with the twin objectives of work and sociability. People seek information for a number of reasons, and may not necessarily start out with a specific aim in mind. However information seeking may be part of a joint enterprise, or may become a way to initiate a joint enterprise of some kind. A number of different studies of work and the artefacts which are used for cooperative work will be reviewed with a view to pointing to just what we need in an instrument which helps us navigate in information space, which will afford the kinds of cooperation which occur naturally in the real world. In the real world, we will find, people do not just do their own work in a vacuum, but rather they do their work with an awareness of other things going on, and a changing level of commitment to joint enterprises.

Navigating 'information space' at present is not like this. Normally query in a library system is different from browsing the shelves. It is mostly done on ones own, and done on a single terminal, which, using time-sharing software and hardware developed twenty years before, interacts with the database holding the details of books and journals extant in that library as if it were the only terminal communicating with that database.

In the real world of books and shelves the reality is different. We look at shelves of books about a subject, and in looking at *these* shelves and not other shelves, we give a number of contextual indicators to others. We can be seen looking at books, specific books concerning specific subjects. Thus others can see our activity and our concerns. Further, contextually, it might even be possible, given a knowledge of schedules, expected essays, etc., to be fairly certain of the position of people in say a degree. What is important is that with this example we can see how peripherally aware we are of the other's concerns and interests, and are able to act on them.

Outline and overview

This chapter will firstly discuss research looking at the practices of work, particularly focussing on how people who are co-present with each other manage to cooperate with others. It will then go on to look at the various ways in which work is accomplished through artefacts, and a few issues which relate to the use and embodiment of artefacts. Both these elements are very useful in thinking about the types of representation and embodiment of people and objects we may require in any type of information 'space' we may design. We will, then, be looking more at the 'social' and also at foundational aspects of any given 'information space' than about navigation in that space. It is felt that the issue of just *what* will constitute the information space is of as much importance.

In looking at the achievements of the workplace, we will be able to see the kinds of requirements for support for a system which aims to support any kind of cooperative venture in navigating information space with the aid of others (for example, engaging in cooperation with others, seeing what others attend to). It will also give us some hints as to the type of functionality for any tools which might aid us to navigate 'information space' of whatever kind. We will look at the types of representations which might be needed in the space, the issues involved in representations of both people and objects, and also some possible ways of instituting an 'information space'. To this end we will look at developments in VR including Cooperative Virtual Environments (CVEs) and other means of representing information.

Work: managing cooperative enterprises in real space

Introduction

Cooperative enterprises of various types make up our work-lives and in a global scale, our society. Anyone who looks closely at a workplace will note the intensely social nature of a lot of work, and just how much people cooperate in joint enterprises in workplaces.

We should not ignore the practices and achievements of those who work in workplaces of many different kinds. All too often these practices are ignored, often glossed over by those who view work as a series of uncontextual tasks, and also often by those who work in the workplace themselves, who often do not fully appreciate the complexity and artfulness of the things that they do everyday, day in day out.

The practices of the workplace are interesting to us in our concerns with navigation of 'information space' because of our focus on and interest in navigation which is personal and social, that is, navigation which is aided by others in various ways. The workplace gives us an idea how joint enterprises are actually achieved in an ongoing way in the real world. In gaining an adequate understanding of this, we can begin to understand the sort of thing which might need to be supported in a 'virtual' way.

It is important to understand that when we discuss 'virtual' navigation in this way we are explicitly *not* outlining the way this may be accomplished. Virtual Reality, one way in which we may explicitly represent an information 'space' in a spatial way, is one of a number of possible ways in which we may accomplish a representation of ourselves and others in a virtual 'space', but it is by no means the only way. There may be many different ways in which we may navigate information spaces without any representation of these spaces in a spatial way.

Prospect

This section, and the next concerning artefacts, will constantly refer to each other. This is necessarily so, as the distinction between work and artefacts is artificial. Work incorporates artefacts in numerous ways, and different artefacts provide various kinds of resources which aid people in getting the work done. The distinction between the two sections is therefore an artificial one. the next section will concentrate specifically on artefacts, whereas this section will concentrate slightly more on communication at work.

Studies of work

Ethnographic and other studies of work and artefacts are very relevant to those of us interested in the idea of social navigation. Our work-lives involve a great deal of information use, and are intensely cooperative. In understanding some of the artful practices which produce the work people do, we can inform ways of working together which may use new technology to let us work with each other in different ways, perhaps over distance, perhaps in different, virtual worlds, perhaps embodied in a different way. To try to aid people working together in an these new ways without looking at this rich canon of work might lead us to some fatuous conclusions of what work involves, which may ignore and gloss over a whole world of the practices, tacit or not, which make up the mundanities of work.

Heath, Luff and Sellen [1] provide a very good summary of some of the contextually rich, mainly ethnographic work done on work which I will quote in full here.

- Cooperative work involves the ongoing and seamless transition between individual and collaborative tasks, where personnel are simultaneously participating in multiple, interrelated activities.

- An individual's ability to contribute to the activities of others and fulfil their own responsibilities relies upon peripheral awareness and monitoring; in this way information can be gleaned from the concurrent activities of others within the "local milieu", and actions and activities can be implicitly coordinated with the emergent tasks of others.

- Much of the interaction through which individuals produce, interpret and coordinate actions and activities within co-present working environments is accomplished using various objects and artefacts, including paper and screen-based documents, telephones and the like. The participants' activities are mediated and rendered visible through these objects and artefacts.

- Both focused and unfocused collaboration is largely accomplished not through direct face-to-face interaction, but through alignment towards the focal area of the activity, such as a document, where individuals coordinate their actions with others through peripheral monitoring of the others involvement in the activity "at hand". For example, much collaboration is undertaken side by side where individuals are continuously sustaining a shared focus on an aspect of a screen or paper-based document, such as a section of an architectural drawing.

- Collaborative work relies upon individuals subtly and continuously adjusting their access to each others' activities to enable them to establish and sustain differential forma of participation in the tasks "at hand". [1]

Work can involve both a degree of detachment from what others are doing as well as peripheral monitoring of what others are doing. In this way we can see in quite sophisticated ways what others are doing. Knowledge of the work being done, and knowledge of the properties of the artefacts being utilized can give us a very great understanding of what others are doing. We can orient to others and to the artefacts we commonly use in ways which could be largely implicit; even explicit communication (for example, a telephone conversation a work-mate is having) can give other work-mates a peripheral awareness of what others are doing. I will now give a number of examples of different work settings to explicate some of the points of these summary points presented by Heath et al.

Robinson, in describing an ethnographic study of work in a nuclear power station [2], gives a good example of how the work of others is understood by an enculturated (someone who is knowledgeable and skilled in the work in question) person, and further, how the inter-relation of people and artefacts in the work is peripherally monitored. Sophisticated understanding of the state of play of the work is attained at a given point before any explicit interaction is initiated.

In the work of the plant, these researchers studied the use of the dials which lined the walls of the control room by the various staff who were responsible for the running of the plant. These dials reported various reading of different parts of the plant, and different aspects of those parts, and aspects of the plant in general.

It was common, these researchers found, for a given member of the control room staff to actually go up to a dial and study it, especially they found if the dial was showing a reading which was regarded as being not quite correct. Moreover, the staff member would sometimes tap on the dial; like a barometer, this might move the dial and change its reading slightly (which might move the pointer into a more acceptable region of the dial). The beauty of this work was that it pointed out that this was *not the only information given by this dial.* There is a lot of implicit information involved as well. At a first pass on this data, one could formalize this interaction in an insensitive way, taking the need to access the dial, and then to check if the dial is reading correctly, and propose that this information could be simply reproduced on individual computer screens which take care of the various elements of the running of the station by the various delineations of staff. they could be seated on individual desks, not needing to move from dial to dial, and perhaps able to check up on other ways if they felt that a dial was misbehaving. However, this is to neglect the *implicit information* inherent not only on these artefacts, but in the relation of personnel to these artefacts and the relations *between* various artefacts.

To explain: we can take the example that one of the engineers goes up to one of the dials, and starts to tap it. An 'enculturated' trained observer can tell a number of things:

1. 'Something is up' concerning the dial.

2. 'Something is up' with whatever the dial represents,

3. This is being attended to by this particular person.

Moreover, any engineer who is familiar with the operations of the nuclear power plant can see an *overview* of different parts of the running of the power plant. This person will have knowledge of the interrelation of different parts of the plant, and so, to take one more simplistic set up, one engineer is tapping at a dial which represents core temperature and another engineer is looking at another dial which represents cooling water pressure, the enculturated work-mate can relate these two things to provide a possible picture of what is going on, while knowing that these two things are being individually attended to. Thus the work is not only done, but seen to be done, it is *played out in public*[1] and *seen to be done* in a context which others can 'read'. With the minimum of effort, others not only perform the work they have to do, but communicate tacitly and contextually to others that they are doing this work. This example has been simplified, however it shows these basic properties.

Heath and Luff [3] give another good example of orienting to others and to the artefacts in the workplace with regard to the work of a control room on the Bakerloo line of the London Underground. They perform a rigorous analysis of video recordings of the work done by members of the control room. In particular, they follow the work of the Controller and the Divisional Information Officer, the DIA.

The work of the two members of the control room is intertwined, yet they each have their own separate tasks. The controller has to make overall decisions about the running of the line, and must sometimes ask trains to stop in the tunnel, or slow down for various reasons. The running of the trains is done to a timetable, which is covered by plastic, on which changes are written. The position of trains is recorded on a large board. Heath and Luff point out that often the DIA will not be explicitly told about a particular alteration to the schedule, but rather will overhear that the controller is changing the schedule, or doing something which will change the schedule. In particular the DIA overhears the controller interacting over the communications system with drivers of particular trains. It works thus; the controller interacts with a particular driver, perhaps asking him to wait in the tunnel. The DIA can see the position of the train from the board, and with the information from the controller's interaction with the train driver (such as a request that the train wait in a tunnel for a minute) this enables him to make an announcement to the passengers at that particular station. This is all done without any direct interaction between the controller and the DIA. As Heath and Luff describe with this exchange[2]:-

```
C:        Control to the train at Charing Cross South Bound, do you
          receive?
(switches monitor to platform)
C:        Control to the train at Charing Cross South Bound, do you
          receive?
Driver:   Two Four O Charing Cross South Bound
C:        Yeah, Two Four O. We've got a little bit of an interval
          behin:nd you. Could you take a couple of minutes in the
          platform for me please?
Driver:   ((........)) Over
C:        Thank you very much Two Four O.
DIA:      Hello and good afternoon Ladies an Gentlemen. Bakerloo
          line information...
```

(Adapted from [3])

At the start of the call [the DIA] scans the real-time display to discern why the Controller might wish to speak to the driver. Even by the second attempt to make contact, the DIA is already moving into a position where he will be able to make an announcement. At the word "couple" he is able to infer exactly what's happening and grabs the microphone to inform the passengers of the delay in the service. [3]

In this case, the DIA and controller work with each other attending and orienting to the various artefacts in the control room, each doing separate tasks yet each able to make sense of what the other is doing with very little direct interaction, rather making sense of the work of the other by using their contextual, situated knowledge of the states of the various artefacts in the room.

'Outlouds', peripheral awareness, and the shift from individual to collaborative activity

It is often the case that people in a work-place may have shifting commitments both to their own work, and to work which is done in tandem with others. The way this comes about will be shown to be interesting to us when addressing those aspects of navigation of information space which are collaborative. Studies of real work-places may again give us insight into those elements of collaboration.

We will now go on to look at the way in which dealers in a City of London dealing room organize their dealing, sometimes alone, sometimes collaborating with each other. There are two things in particular we will note: first, the initial signalling of the 'bid' by a particular institution and second, the collaboration which is initiated. This work is by Heath and his collaborators again [4].

We can take the following exchange which they report (below). To set the scene:- Anne is across the busy dealing room from Robert and John. Robert and John are dealers who sit next to each other. All the dealers sit in rows, each with a number of screens on which information is displayed. There are telephones and 'stentaphones' (speaker-phones) constantly in use, so the ambient noise level is high. John and Robert occupy the 'number one' and 'number two' positions in that dealing room (i.e. they are the most important dealers in that dealing room).

Anne:	HAN:SON. TWENTY OF AN EIGHTH, FORTY BY FIFTEEN:, (SHEARSON) ON THE BI:D [Shouting the information[3]]
Robert	((Sitting down))
	(0.2)
John	Are we going to hit 'em?
	(2.3)
Robert	((Peers at screen))
Robert	Erm::^ (.) YES::,
Robert	(0.9) WHO'S THAT? [to Annie]
	(1.0)
Robert	Bernie? ((Picks up phone))
	(3.0)
Robert	We want to sell (forty)...

In this exchange, Anne shouts across the room that Shearson (a large broker) are 'on the bid' for (wanting to buy) Hanson shares. Robert is in the process of sitting down and gives no sign that he has heard her. As Robert sits, John turns in his direction and asks whether they should "hit them", i.e. sell them Hanson shares. As Heath et al. point out:

John's question assumes that Robert has both heard the utterance and may be prepared to collaborate in selling the stock. The utterance appeals to, and invokes a mutual orientation towards, selling stock and initiates collaboration with John.

Robert then lifts the telephone, and also shouts to Annie asking who is on the phone to her. In minutes, a large amount of stock is sold.

Heath et al. discuss the shouting as follows:

Whilst Annie's utterance is shouted aloud, as if potentially relevant for anyone within the local vicinity, it is perhaps only two or three traders who might have an interest in the information. Shouting out loud, rather than specifically telling certain colleagues, is not only a relatively economic way of informing a number of recipients, but also delivers the information in a way that does not necessarily demand that anyone responds.

As we said earlier, there are two things of interest here. I will start with the collaboration first. Here we see two people who mainly work alone but who do collaborate as well. We see the way in which such collaboration is initiated in this setting by seeing that the other is available, knowing practically that the other is in possession of the same information. Seeing the other's work in context and seeing the other's awareness of the same concerns helps the initiation of collaboration.

The 'outloud' is of interest in its own sake cf. [5] as a way of economically giving information to a large group of people. It is of interest here as a way in which people keep peripherally aware of goings on and can invoke "a mutual orientation towards selling stock and [initiate] collaboration" [4].

Relevance of this work for Social Navigation

Let us go back to what we discussed earlier about personal and social navigation of information space, and draw in some of the issues discussed above. So far we have looked closely at the way people work in a number of settings where they are co-present. We have seen a number of things about the way they work which helps them work in these settings, which have been uncovered by careful ethnographic work. We can begin to apply this work to navigating of information spaces when we begin to think about the elements of this activity, and the contexts in which these elements exist.

We have seen how people work together in the real world situation of the office. People attend to their various tasks, while maintaining awareness of what others are doing, and they may join others in a joint collaborative task, as well as work on their own concerns.

This is important when we think about ideas of jointly navigating information space; 'asking the other the way'. To take some of these real world abilities to the world of information starts to point to some basic requirements of the *information space itself*, never mind the navigational tools which are produced to aid navigation in the space.

Seeing what others attend to in a context

The most obvious point to come out of this work is that we have to see just what others are attending to and what they are doing. If we wish to collaborate with them or ask them the way in an information space it is useful for us to know such things as what their interest is, what they are doing, just how busy and engaged they are. In the study of the city we discussed earlier [4], it was found that potential collaborators would orient to and act on activities that signified that the activity which was taking up the other's attention was over. For example, a potential collaborator on a deal in the City of London dealing room waits for the other to finish writing out a deal slip (signifying that that particular activity is nearing an end). We can take the lesson of this to a putative virtual world where one is able to engage others to help one in searching for information. One must have some possibility of seeing:

1. that the other is available (or conversely that the other is occupied),

2. that the other has similar concerns.

As has been argued by other researchers (cf.[6]), we must remember that what we are trying to support is the *achievements* of those actions in the real world and not the practices. However, conversely, we must be careful that it is these achievements which we are supporting. And that we have correctly deduced what the achievements actually *are*. This is one reason why one must take such a close look at *the context in which these practices are performed.*

It may be that the information navigation is instantiated in some sort of VR representation of an information space which involves representation of the various elements of knowledge in that space. This may be in terms of books, journals, papers, or even web pages. There are already a number of systems of this type which will be focussed on in a further section of this chapter. The work we have described before suggests to us that the ability to see what people in the space are attending to and further what they are doing with the information is worthwhile. On a broader level, the research above hints that our ability and wish to collaborate is tempered by the needs we have to do our own tasks. Collaboration has to fit in to what we are doing. We come into the need to represent engagement and activity so that basic politenesses can be observed. This is especially true in an environment where everyone will synchronously inhabit the space (as in OPAC querying in a library system, where we may be searching for books in a particular field and see that others are there as well).

Scalability and representation

A further topic which has not yet been considered is that work-place environments which we have detailed above are necessarily *limited* spaces. Our virtual information space may be one which spans the globe. Thus it will be necessary to attend to issues of scalability. We will have to contend to the possibility that in virtual form, many more people may be able to attend the information spaces that are created, than would be able to in a real work-place. We must have possibilities of representing these people and this space which take into account this.

In the next section on artefacts, we will begin to deal with a few of the issues concerning representation. We will go on to explicitly look at the creation of different types of VR spaces (and particularly the creation of types of multi-user spaces) by various research groups, and note the possibilities for representing information and information spaces that this research shows us.

Virtual worlds: contexts for interaction

We have also seen how in the work-places we which have been discussed that people largely interact with attention to artefacts, rather than face to face. Such is our belief in face to face interaction being of primacy in situations of co-present activity, that we tend to call a co-present meeting a face to face on. The rôle of artefacts is overlooked. However the exposition of these studies shows, hopefully, how we work with and through artefacts. We also monitor each other peripherally, overhearing and 'looking over their shoulder' as we work.

Perhaps in the domains where we are navigating different information spaces we will not be party to such a close set of ties as we might observe in a work situation. If we do not collaborate, no aeroplane will crash, no bank will lose millions, or no underground train will crash. However, we take it that there may well be good reasons to collaborate. We may gain knowledge of domains which we know little about, by following, or being shown the way by, experts. We may be able to reciprocate by aiding others in areas where we are expert. How exactly this would work cannot be predicted, and to have any idea of this we may have to look at the foundations of society itself to give us an idea. However, to aid collaboration, then, we have to be able to see what the other is attending to, and have some 'common artefact' which we can share. In this case, the common artefact is likely to be some kind of representation of the knowledge object.

To pursue this line of thought further, we will now look at the use of artefacts per se in the work-place. Of course, the section following and the one above show a very false divide. We have already seen some artefactual use for example in the studies of the underground. The real time display and the time table and video monitors give a context for the DIA to understand what the controller is doing. The next section, however, will bring the artefact, and particular issues which concern it, to the fore.

Artefacts

An essential element to the everyday embodied environment of the workplace is the artefact. This can be a tool a screen, a chart amongst a number of possibilities. The essential importance of artefacts is that they can, in various ways, embody what we know and regard about the work done in the work-place, and do so in ways which are able to be read by others. They can further embody philosophies, and politics. They are essential to the way we work in the workplace and in the world by providing different types of repositories of information, and ways of sorting and sharing knowledge. Some of these benefits may be to provide explicit or implicit ways in which to communicate knowledge. Many aspects of work are done in relation to artefacts.

Embodiments of the state of work

Artefacts can provide ways in which they embody the state of work in a number of ways. This embodiment of work must necessarily take place in relation to a set of developed practices. Thus, artefacts are able to embody information in the workplace and communicate it to others. This will be illustrated with a number of examples, looking at different aspects of artefacts in use, in a number of different work places. An essential point to take from these examples is that artefacts in use are products of and necessary elements of, the achievements of the work done in these workplaces.

Flight strips in an air traffic control room

A group of researchers mainly from Lancaster University did a long and detailed ethnographic study of the work done at an air traffic control room (e.g. [7]). An important aspect of their findings for this work is the utility of the artefacts which represent flights in the work of the ATC staff.

The flight strip is card, about 8 inches by one inch, which is divided into fields and which contains information for one particular flight. It includes the call sign, flight level, heading, air-speed, planned flight path, navigation points on the route, and estimated time of arrival (ETA). Strips are printed off out of a database which holds the flight plan which has been filled in by the pilot of that particular flight prior to the flight, and which then is modified by radar data and sometimes by inputs keyed in by controllers and assistants as the flight proceeds[4].

Next reporting point | Current a/c orientation | Type/make information | A/c squawk identifier | Route information | Previous reporting point

The strip, Hughes et al. argue, becomes "more than a repository of information". Rather, the strip becomes a "work site". The strip becomes a device for recording the various elements of the flight. These elements are not recorded any old way. there are specific conventions for the way in which these elements are recorded. For instance, when an instruction is given to a pilot to ascend to flight level 220, s/he marks this on the strip: in this case, with an upwards arrow and the number 220. When the pilot acknowledges the instruction, the controller crosses through the old flight level on the strip. When the new height is attained, the controller makes a check beside it. [7]

The strip becomes a memory of the flight. We can see from the above that each stage has a formalized procedure which makes the position and progress of that flight readily available. It is important to note that the strip is more than a memory; it is also a *public* memory and a resource for everyone working in the control room. The strips are used in conjunction with radar in the work of the ATC staff, in particular ways: The strips are used more for planning. As Hughes et al. say:

The strip acts as a notepad and work-site for all the team members, any of whom may write on it, and each of whom uses a pen of a different colour so the source of the annotations is immediately apparent. Anyone who notices a problem with a pair of strips- perhaps two flights at the same navigation and at the same height- can 'cock out' the strips'- move them noticeably out of alignment in the racks. This makes it immediately obvious that, when it becomes time to deal with those flights, a problem will need to be considered. [7]

Hughes et al point out that this means that changes in direction, speed, ETA etc. are not lost. Because they are annotated directly on to the strip a lot of information is available to the reader. A reader has an idea that there may be a problem (types of annotation, numbers of them) who addressed these issues, and has an idea 'at a glance' of the state of the work (by the amount of 'cocked' strips there are on the racks, for example. Thus the strip not only serves as a means of giving flight details, it is as much a place where work is both done and documented, and able to be traced, with overall representations of the state of play. Workers have in the strip a resource to give *peripheral awareness* of other's actions.

The strip is a potent artefact; it is a vital instrument for ATC. It is possible and on occasion necessary, to perform ATC with the strips but without radar. To do so with radar but no strips, however, would pose extreme difficulties. [7]

Portability between different populations: Boundary objects

Sometimes, an artefact can embody aspects of many different populations of workers (or 'communities of practice') with many different concerns. Leigh Star and colleagues discuss artefacts of this type with reference to Latour.

A boundary object, argues Star, can be typified by a map. A map is a representation of space, but a representation of a particular type. It gives various types of details of the landscape, and representations of various elements. It is and object which can be used by different disciplines in different ways (e.g. biologists, hydrologists, archeologists, farmers) and each discipline will take particular features from the object. However, the representation of the landscape, though it is used differently by the different groups, essentially stays the *same*. Star uses the term 'boundary object' to reflect the observation that these objects stand at the boundary between disciplines (or different 'communities of practice'). Each community, whether it is hydrologists and biologists, orients to the object in a different way, but each community can use the same object. These objects can become a way of communicating between different communities, and so between different levels in systems, different processes in manufacturing.

The boundary object seems a necessarily vague thing, including in it multiple representations giving the possibility that different communities take a partial view of the object away with them. Thus two or more representations can overlap within the same object. Also, even if each group takes the same types of data from the boundary object, they may use the information in a different way.

Embodiments of politics

As well as being the embodiments of work, artefacts can as well be embodiments of particular world views, or particular types of politics. A canonical example of this is with regard to a bridges built on Long Island, known as the Long Island Bridges [8]. These bridges, designed by Robert Moses clear the curbs below them at just over nine feet height. Thus traffic passing on the roads under the bridges have to be less than nine feet in height to pass. New York public transport in the form of buses has a height of twelve feet. Thus Winner argues that these bridges effectively exclude those groups for whom public transport is the only possibility of travel, favouring the car-owning middle classes Bowers [9] takes this argument slightly further. Winner has evidence of Moses' political views and affiliations. However, they are not necessary for their political impact.

The bridges have political consequences whether these were intended by the designer or not. Indeed, the politics of the Long Island bridges depends more on the relation between two heights (nine feet and twelve feet than Moses's well-attested conservatism. [9]

Bowers is making an interesting and radical distinction here, between what we may intend to happen with our designed artefacts and the political ramifications, which may be intended or may be not. It does not matter what Moses intended. The ramifications of the bridge heights, whether intended or no still had the ramifications of excluding those who were solely dependent on public transport (largely working class and a disproportionate amount of the largely working class black population).

This particular case demonstrates some possibilities for conceptualising aspects of the built environment. However as demonstrated, the design of artefacts can have a political impact at a much more subtle level.

Caveat: artefacts and representation of work

We must be very careful to determine *just what the artefact can represent.* Take the example of an order form. This could be taken to represent the different stages in the work of a particular workplace. However, if it was embodied computationally, this computational artefact may force on the workforce a particular set of working arrangements, for example enforcing that everything must be done according to a set sequence. Button and Harper [10], studying a furniture factory, looked at the impact of such a computer form system, which replaced one which was a paper one. This system had been created after some effort by the system developers to study the work stages of the company. In their system, these stages were embodied in the program; it enforced a particular order, one that the system designers had seen in their study, in which an order was processed in a particular way. A piece of goods would be ordered in the front office, where a price was sorted out, and then would move through the various stages of the factory.

However, this model of the production not actually accurate. Button and Harper found that established customers would phone directly to the first manufacturing stage, and check the availability of the raw materials, time it could be done in etc. The order would be in progress long before any attendant paperwork was. It was often later, and post hoc that a price would be sorted out and the paperwork would be done. The old system, of course, allowed this. Forms were often made up (for established clients) in a way which deviated significantly from the ideal but which was fast, and simple to do.

What we see is that the new software could be seen to:

1. actually have the capacity to enforce a certain order of production (by making it difficult to change the order and thus for the cutting shop to accept orders),

2. enforce rigidity in a way of working which was ad hoc and responsive.

Thus we can see with regard to this example that the artefact (the order form) did not actually represent the chronological stages of production, but what they usually were, and what was oriented to in the ad hoc working practices of those in the shop floor (i.e. in the 'working up' of the order form later). The system designers took the form to describe the working practices, and designed accordingly. The design which was meant to aid the processes of the factory could actually stand in the way of them.

Implications of work on artefacts for social and personal navigation

We have seen in the previous work how artefacts in the workplace have an intense relationship with the work itself. In fact, so close is the relationship that often the work cannot be done without these things. Work is, from the examples above, intensely artefactual. However, we have shown further that while artefacts of various types (such as timetables, representations of train-lines, air traffic control flight strips, etc.) Can give representations of the state of the work, they can afford generalizations in various ways, giving the enculturated user (one who is fully aware of all aspects of the work) overview of the work, "at a glance" awareness of the state of different aspects of the work.

Some type of representation of information in an information space may have a great deal of utility if they can capture some aspects of the utility of these things. To be able to see "at a glance" the state of a search, where one has been and where one wants to go, to be able to see the kinds of searching possible and the ways one can go. We might also see what others attend to in their searches.

We have also shown how certain types of artefact can be portable across domains of knowledge, providing a conceptual bridge between quite different knowledge domains and specialisms (or "communities of practice", ref.). Portability is one aspect of certain types of artefact discussed by Star amongst others. In a navigational environment it is likely that searchers (or navigators) will have different needs, and different outlooks. A class of these 'boundary objects' will surely be a part of any scheme of things for a truly universal representation of knowledge.

However, we have also come across caveats to the utility of artefacts, however they are defined, whether they are simple representations of different types, or actual mechanisms, whether they are paper or computational. We have seen Winner's example of the Long Island Bridges the overt political outcome of certain aspects of technics, in this case a bridge. In this case we can see from Winner's work that there is a very strong case for this outcome being highly intentional. However, this is not necessary. More worryingly, artefacts of various kinds can have an overt political impact whether or not this impact was desired (or even explicitly abjured) by the designers. The example of the Long Island Bridges is a particularly clear cut and obvious example. The outcome of other types of artefact may be less overt, but still as pernicious. Sproull and Keisler [11] could be seen to expand and extemporize on this theme with their discussion of first and second level impacts of technology. There may be a basic impact in terms of utility of one form or another, but there may also be dramatic and unforeseen secondary (usually social) impacts. Postman [12], for example argues forcefully, using methodology which goes back to Mcluhan and further to Lewis Mumford, cf. [13] that the cultural impact of the television is far beyond that of providing entertainment and news from around the country (and eventually around the world). It has also, he argues, given us a new way of conceptualizing the world, in the form of entertainment. Postman argues that this conceptualization has directly lead to the absence of any substance to political discourse, with a reliance on appearance rather than substance (for appearances are the one thing that television excels at transmitting).

We can go further from the ideas that artefacts can have embedded in themselves:-

1. representations of the progress of work,
2. embodiments of work,
3. political intent, whether meant or not.

ARTEFACTS CAN HIGHLIGHT ASPECTS OF THE WORK, AND ASPECTS OF INFORMATION WE MAY NEED IN OUR WORK. HOWEVER, THEY CAN ALSO FREEZE ASPECTS OF WORK WHICH WE DO NOT WISH TO DEAL WITH. A KEYRACK IN ITS CONSTRUCTION FREEZES A CERTAIN RELATIONSHIP OF ROOMS, FOR EXAMPLE. AN ORDER FORM FREEZES CERTAIN RELATIONSHIPS OF CONCEPTS. HOWEVER, THIS FREEZING CAN GIVE RISE TO POLITICAL ISSUES OF REPRESENTATION. WHEN WE CATEGORIZE A BOOK, FOR EXAMPLE WE TAKE SOME ELEMENTS OF THE BOOK AND DISCARD OTHERS. WE DO NOT TAKE THE BOOK'S SIZE, OR APPEARANCE (WHETHER SOME AESTHETIC JUDGEMENT OR NOT), NEITHER DO WE TAKE SOME ESTIMATION OF THE BOOK'S QUALITY. THIS IS NOT TO SAY THAT WE SHOULD, BUT RATHER TO HIGHLIGHT SOME VERY CLEAR ISSUES FOR THE POLITICS OF REPRESENTATION WHICH WILL AFFECT DIRECTLY ANY KIND OF NAVIGATION IN AN INFORMATIONAL SPACE. HOW THE INFORMATION IS REPRESENTED IMPINGES ON JUST WHAT IS INCLUDED IN THE REPRESENTATION AND WHAT IS LEFT OUT.

REPRESENTATIONS HAVE OF COURSE THE POTENTIAL TO GIVE US A WHOLE RANGE OF POSSIBILITIES OF HOW TO NAVIGATE IN A VIRTUAL SPACE OF INFORMATION, RELEASING US FROM RELIANCE ON DATABASE QUERIES. HOWEVER, THE WAY WE NAVIGATE MIGHT WELL BE RADICALLY AFFECTED BY JUST WHAT WE MEAN BY REPRESENTATION.

Virtual Spaces

Introduction

Elements of the field of virtual reality or VR have begun to have an impact on areas of CSCW, especially with regard to conferencing and other types of work through computers. The focus of this work is enabling a representation of a person to be able to work with other representations in much the same way as one can do so in real life. It aims to provide people who work together with something like the kind of "natural social conventions" [14] which we use in the real world of work. Another element of this work is being able to provide people with the kinds of "mutual awareness of each others' activity [which is] essential to the coordination of interaction in the workplace and the smooth running of cooperative activities" [15]. We have seen from earlier work mentioned above, that co-present work is usually intensely *artefactual* in that orientation is usually to artefacts in the workplace than directly to other people, orientation to other people is usually done in some kind of peripheral way, where workers peripherally monitor others in such aspects as how busy they are. If they can be interrupted, invited to take part in a cooperative venture [4], what are they attending to, and what they are doing. It is thought that VR technology could go far beyond the kind of abilities of representing the other as was possible by the kinds of video conferencing systems which have been so much talked about and researched in the short history of CSCW.

This technology has two particular applications, both are of interest to work on social navigation. One, we have already mentioned above, is in terms of the abilities of VR to be of use in rendering representations of some kind of humans inhabiting the virtual space, and of aiding them in co-present meetings. The other application is in terms of aiding humans in retrieving information, and in particular, aiding humans in retrieving information aided by others, and by the representations of information that the system creates. This may be in the form of information terrains, where information in a dataset of some kind is represented spatially, and where the user can 'fly' over the terrain until she or he finds what he or she wants, or which provides an 'overview' of the field of information, which may show things that another kind of overview cannot.

A particular application of virtual reality work has emerged in the last few years. A lot of this work comes under the term of CVE, standing for Cooperative Virtual Environment. This work is notably different from common notions of VR in that it is usually not immersive, but rather gives a series of spatial cues represented on the standard 2D computer screen. In effect, this kind of virtual space is a 2 _ D one, in that it provides information to the user in the form of motion parallax, etc. Like games such as Doom, this can give a sense of immersion in the virtual world whilst requiring none of the specialist headsets, data gloves, etc. Motion is accomplished using normal keyboard and mouse commands.

There is an additional point to this. Immersive VR is by definition one where the aim is to immerse the user in another world, and concomitantly, detach the user from his or her immediate world. This may be useful for certain types of activity, such as simulation, however it is not useful if the user has to monitor the world around them. As said previously, a characteristic of real work is that people take part in multiple interleaved collaborative ventures, sometimes cooperating very closely with another, at other times more detached. However there is constant monitoring of others and of artefacts and activities in the work environment. Our work environment is a set of *mixed realities* of various different virtual worlds, information spaces, artefacts, communication devices. Thus immersion would isolate us from our immediate information space and some of the array of other realities apart from the virtual one.

Collaborative Virtual Environments

CVEs such as MASSIVE and DIVE have been proposed as ways in which people will be able to work together in a virtual space which may go beyond more 'traditional' CSCW technologies such as media spaces (RAVE, etc.) or other teleconferencing and multimedia technologies (such as BT's VC8000 system,[16]). The CVEs are proposed to enable people to be present in a virtual environment, and work together. As Benford and Fahlén argue [14], previous CSCW conferencing systems incorporate a range of mechanisms for floor control such as chair people, reservations and token –passing. However, they also argue that work-flow techniques (e.g. Coordinator; [17]) in general also represent a form of "conversation management". They argue:-

... we believe that these approaches are generally too rigid and unnatural to be applied to spatial settings. As an example, a real-world implementation of explicit floor control would be tantamount to gagging everyone at a meeting and then allowing them to speak only by removing the gags at specific times...

New techniques are needed which support natural social conventions for managing interactions. One approach might be to take advantage of the highly fluid and dynamic nature of space. [14]

Thus CVEs seem to offer the possibility to use more everyday interactional competencies. However, they further offer the possibilities of utilising the kinds of mutual awareness of work activity, and interaction with suitably embodied artefacts, which I have talked about earlier with regard to ethnographic studies of work.

Some themes of interest relating to CVEs and other types of virtual worlds

Accountability: social production of everyday interaction

In Studies in Ethnomethodology, Garfinkel [18] discusses the ways in which we can conceptualize how people obey social rules. Rather than adopting the perspective that other theorists such as Parsons took, that social rules are internalized, Garfinkel takes a radically different perspective. As Heritage argues, Garfinkel discovered, through breaching experiments, that the constitutive expectancies of the attitude of everyday life are treated by mundane actors as profoundly normative, and morally sanctionable, matters. [19]

Taking the simple example of a greeting, he talks about the choices that someone faced with a greeting will have. One can respond to the greeting, or not respond. To not respond, or even take too long to respond, takes the action away from the realms of the 'life-as-usual' stance, to one where an account may be sought by the other for his/her breach of the norms. As Garfinkel would argue:-

With respect to the production of normatively appropriate conduct, all that is required is that the actors have, and attribute to one another, a reflexive awareness of the normative accountability of their actions... normative accountability is the 'grid' by reference to which *whatever* is done will become visible and assessable. [19]

We are not automata, Garfinkel argues, even though at times we react to a greeting without a great deal of thought. There are times, given this knowledge of the norms one can use this to *refrain from replying*, and so set in motion a search for an account (e.g. you were really drunk last night and spilled wine on me). This is only possible because this knowledge of norms is *reflexive*, that is, it is knowledge that others have this knowledge of our knowledge of the norms, and so on to infinity. W may get the wrong idea from this. We do not do this kind of sophisticated reasoning all the time, however, we can use these norms in this kind of artful way, and do. Garfinkel argues that we are not "judgmental dopes", but neither are we (at least mostly) calculating machiavellians.

Now the interesting element of this with relation to a discussion of CVEs is the *sort of accounts* able to be generated. In the real world, knowing that the person is not deaf, and is in range enough for the greeting, is usually be enough to determine that an account, such as "I am in their bad books" may be required. Notice how everyday life allows us to quickly leap from a technical account (such as hearing) to an account which is rooted in the non-reply being *meaningful*, e.g. that the other person is displeased with one. However, in the case of CVEs, this search for account is somewhat increased. A significant part of this is because our embodiments in a CVE, that is, our shape, position, ability to talk, are all *mediated by technology*:- servers, computer networks, intervening layers of software. Thus the search for an account becomes ever wider, and it is much more difficult to narrow down accounts.

The virtual world is still part of the real world

Conventional CSCW wisdom sees that we can interact in CVEs much like we do in the real world; we have affordances of space, movement, communication and embodiment, as we would have in real life. But, as Bowers et al. argue, what we have forgotten here is the effort required to achieve all this. If one 'meets' someone in a CVE space, it is not at all easy to tell what is happening 'in the breach'. There are too many possible accounts to be made: the server is down, there are technical problems, etc. In a person used to the virtual world, there is a cornucopia of possible explanations. That is because what has been forgotten is that the CVE is very much part of the real world, as are the people embodied in it. Bowers et al point to a number of issues which arise in CVEs as a result of their careful analysis of CVE encounters, particularly with MASSIVE:

... in the introduction, we noted the central importance of embodying users... are the embodiments trustworthy as bearers of meaning and significance?

Yes, indeed, they can be. *If* activities in the real world are adequately aligned with the embodiment's activities in the virtual world, *if*- for the purposes at hand, the embodiment adequately displays its attentiveness to ongoing activity, *if* participants are adequately mutually focused on a common thread of activity, and *if* the machines are working, *then* the embodiments can be relied upon both by their 'owners' and by the others trustworthy resources for social interaction. [15]

'Achieving' a conference

Embodiments are not real. They are an achievement of programming and the concious effort of people partaking in the interaction. They are figments of what happens in the real world. However, what happens in the real world can sometimes disrupt what happens in the virtual world. Bowers et al. report the term "corpsing" as in to "corpse" for when an avatar stops moving or interacting with the virtual environment in any way, as a result of technical problems, or as a result of something happening in the real world which has distracted the person represented by the avatar, e.g. a telephone demanding the attention of one of the members. They also describe cases where *real world* interactions can be heard in the *virtual world* and understood in a *completely different context* from the one in which they were said, for example, Bowers et al relate the following example: in the real world, a person co-present to the 'owner' of an avatar has moved the avatar, and the person whose avatar has been moved asks the other "What are you doing?". This is all perfectly understandable when viewed in this context. However, in the *virtual* world, what is experienced is that an avatar moves, and then that avatar asks the *other avatars*: "What are you doing?". An account for this is then given by the 'owner' of that avatar, explaining what has happened.

Thus, in the CVE, there is still a short-fall in terms of the mutually understood reality of users (cf. Schutz [20]) In the real world, we do not expect that we will suddenly be faced with problems which are the product of the breakdown of reality, the sort of problems we only usually face when partaking of extremely psychoactive substances ("Ugh, you're turning into a slug!" "No I'm not!").

However, Bowers et al go on to argue that nevertheless people are able to repair these problems and work together in virtual worlds. Their work helps to uncover the artful work required of the participant in order to take part in these kinds of meetings:

We see researchers and meeting participants engaged in their day-to-day business, exploiting local knowledge, utilising everyday skills and competencies, artfully managing contingencies and problems as and when they come up as best they can. We see ordinary interactional competencies (methods for managing turn taking, displaying attentiveness and orienting bodies, using other means if one fails) deployed so as to make less familiar, less trustable arenas for action more recognisable and reliable than they otherwise might be.

Mixed realities

As said earlier, there has been a general interest in explicitly acknowledging the role of the real world in the virtual encounter. This is one of the elements of the concept of *mixed realities* where more than one type of reality is present in the system. It is partly conceptual, as all technologies of this type have embodied one or more 'realities'; that of the real world and the virtual. Earlier work on video conferencing also alluded to this kind of issue, especially Dourish and Bellotti's work on virtual office sharing. A particularly radical step in this direction is the *Internet Foyer* [21] which incorporates a number of different types of reality (and application).

The goal of this particular work was to try and replicate some of the elements which make a real foyer so useful; as the public face of an organisation, as a point of entry, as a place where one can find help on the organisation, as a place to sign in, as a place where one can rendezvous with others in the organisation. They start by conceptualizing WWW 'home' pages as a 'foyer' (as they do incorporate some of the elements of representing an organization). However, as they argue, WWW pages have some major flaws when compared to a real foyer:

the people who pass through them are not generally visible to one another or other observers (a major criticism of current WWW technology in general). Thus, security may be compromised and there are no opportunities for rendezvous and social encounters. [21]

The Internet Foyer is an attempt to merge a virtual foyer with a real world foyer and provide an integrated shared space which spans both the physical and virtual worlds. Thus, visitors entering the organisation's WWW space would be able to interact with those entering its physical space and vice versa. [21]

Thus a number of elements make up the Internet Foyer. There is a real physical foyer, there is a CVE which is based on the DIVE toolkit, and there is a visualization of a number of nodes on the WWW which have particular relevance to that organization. The visualization is produced by an application which produces a visualization of a set group of WWW which group strongly interlinked nodes into spatial clusters. Another application can chart arbitrary regions of the Web and represent nodes by the number of links going to them; pages with less links are represented as lighter nodes, and strongly linked pages as heavier ones. Thus what is produced is a 3D (actually 2_D) representation of a group of WWW nodes embodying a fairly sophisticated amount of information.

The Internet Foyer thus enables a number of different types of visitor to converge on it in a number of different ways. A physically present person will see a graphical visualization of the virtual foyer. This also gives representations of virtual users of the foyer via avatars which move round the nodes of the visualization. Also shown are less detailed avatars which are representations of users currently accessing the set of web pages. This last group see only a set of Web pages, as their browser allows.

Virtual users of the system which are accessing it as a CVE see the same as those in the physical foyer, but are able to navigate around and go to specific objects which may be others or WWW pages, when an object is selected, a WWW browser is launched which displays the particular web page. CVE users can communicate with other CVE users and those physically present by means of an audio channel, and can also see those in the physical foyer by means of a real-time video window which is mapped onto one wall of the virtual foyer.

This system thus incorporates a number of types of mixed realities, and attempts to provide something like a seamless mix of different realities which has some echoes of the work of Ishii and his collaborators on video-conferencing systems e.g. [22] Whether it is successful in this is not currently known. However, for current ideas of social navigation this work in particular has some interesting possibilities.

Discussion of the work on VR

We have looked at a number of possibilities for working together that virtual reality technology may provide. It is still at the stage of being largely experimental. However, at this stage some technologies capable of embodying people in something approaching a virtual world are available. Such things as the Palace, Onlive! Traveler and some others are of some impact. It is still unknown how much. Bowers et al, in their early attempts to look at virtual interaction from the rubric of interaction (and conversation) analysis, makes a good start in looking at the virtual encounter as an achieved artful entity, one that is effortful and needing of some skill. As they say themselves, this is early work, and there is still a moot point concerning just how much relevance these observations and findings are for other virtual interaction (such as collaborative information retrieval). However, some points about the work discussed above seem to remain general. We are not our avatars, and never fully will be. We are part of a real world, with those particular loyalties and ongoing projects which go on in it. Our connection with our representations seems necessarily contingent. Perhaps this itself may provide the foundations for new ways of interacting. Like videoconferencing, it may be that new modes of interacting appear, and that new social norms grow out of experience in the virtual world.

Social navigation, work, artefacts and virtual worlds: discussion and possible conclusions

We have shown, in recounting and discussing various studies of work, the ways in which people take part in work; monitoring others, and seeing in a context the activities of others who work alongside. We have seen the ways in which people can gain an overview of the work, seeing, with the use of artefacts and the knowledge of others' activities, the state of the work at any given moment. Further to this, we have shown examples of the ways in which people shift through out their work-lives from individual activities to joint ones, and the kind of peripheral awareness, and knowledge of the context of others' activities which is needed for this ability to initiate collaboration without upsetting the progress of others' work.

As well as discussing the intensely artefactual nature of work, we went on to look at studies of artefacts used in the context of peoples' working activities. Artefacts can do a number of things. They can embody different elements of, or an overview of work, give clues to the way in which work is done, and may be ported between different specialisms, or 'communities of practice', sharing knowledge between different fields, as each specialism looks on the object in a different way. However we have also seen how artefacts can have political implications, and can embody politics, whether meant or not. The work of Button and Harper [10] shows us how an artefact can be seen to portray the work in a way in which it actually does not, in this case positing a sequential order for work which is not, but rather is ad hoc and 'worked up' later. We must be careful to make sure that we pay attention the achievements of work, rather than work practices. However, we need to look closely at work to make sure we have identified correctly what these achievements *are*.

Some directions from the research studied

How then, do we take these achievements which have been described in the various studies of work and use them in a way which may apply to navigation of information space? We can think of a number of aspects of this work which may apply.

We need to see that the other is attending to something that holds our interest. If this is not in the real world, and we are not looking at the 'personal and social navigation of information space' simply as another set of leaden metaphors which burden our perception of the working environment, then we are necessarily seeking to look for a means of representation for three things: the person himself or herself, and the domain in question, that is, the information space itself, and the informational objects which are situated in this space. We need representations of the three because we are not just interested in one person's view of a particular domain of information, but rather, we are interested in cooperation in a posited information space. The further element to people and the artefacts (the informational objects themselves) is not only do we need a representation of people and the artefacts which they may use within that space, but further we need some way of situating people and artefacts in it.

From the work on cooperation in real-world working environments we went on to look at virtual environments and collaboration in them. In a virtual environment, it was pointed out, one has the possibility of representing something like the amount of information which we would have in a real 3D environment. In particular, we may gain something like the contextual understanding of others' activities and intent that we would have in a real-world setting. However we must be careful to incorporate in our thinking of possible virtual representation of people and artefacts, the issues of representation which we have talked about with regard to artefacts: they can embody politics, and they can enforce certain ways of working on people which may disrupt ways of working which have developed over time. In conclusion, there are a number of issues which we must bear in mind when we attend to just how we can represent information and information seekers in an 'information space'.

This may not just be an issue of representing to what people attend. There is the further issue of transcending the limits of the work-place. Of course, we do this everyday in a number of ways, with email, video conferencing, the telephone, even letters. The difference between these technologies and a VR technology is that we have the *possibility* of bringing the real world skills of the office into a means of representing information (should we find a means of doing so which is satisfactory). Further, this representation can *overlay other media*, especially in the case of Web use, or accessing a particular newsgroup, or searching a database, such as a library. Thus there is the possibility of finding and interacting with people with similar concerns, in the case of the internet, in a way which scales down to fine distinctions of interest.

Representational issues

As we said above, the process of representation is not straightforward. Let us think of an object in a virtual world. Objects in virtual worlds are representations of various kinds. We can see this in various VR systems: recent work [23] has shown the possibilities of using representations to tell us a lot about who we are interacting with. Rather than the usual avatars, which generally represent one person (but which may represent more, a situation which may mislead and is worthy of research,) there are larger avatars which represent a group to someone of sufficient distance away. as we come closer, the large 'blockie' changes and one sees a load of smaller blockies. This is one way of coping with scalability problems (when instead of the standard four or five avatars, one may encounter hundreds). Here is a way in which a representation of the avatar has been altered because of something systematic happening with the avatar. It is also a way in which this information can inform others. Therefore here is a way in which the representation of an avatar can inform people of a property in the thing (e.g. a person) which the avatar represents.

Let us however think of a library. With regard to the books in the library, we have available to a reader the physical *embodiment* of the book: its size, cover details, etc., right down to the typeface on the book which may give an indication of the *genre* of the book. This is separate from the positioning of the book, is another rich source of information about the book, for example, a textbook usually is of a particular size and thickness, with a particular typestyle. We can easily and in a sophisticated way, differentiate between a medical journal and 'Hello' magazine. In this example, as in the example of the 'blockie', the physical aspects of the object give much meaning. It is not inconceivable that some of this richness could be introduced into representing objects in VR spaces.

This is not the only way in which information can be represented in virtual space. We have the capacity to represent information not only by the avatar, or virtual object, but also by the placing of the object according to some sort of scheme. The virtual space has three sets of coordinates (one of these is of course 'virtual' as the space is represented by a 2D view, but made 2_D by a mixture of perspective and motion parallax). We can represent information in a virtual space in a way in which the dimensions actually mean something. Some researchers have done this, representing sets of objects, such as books, where a conceptual closeness is represented by a physical closeness in the virtual space.

One can theoretically conceive of a virtual information representation where each of the axes, x, y, and z has a meaning, giving those who understood this a means of seeing 'at a glance' a lot of knowledge about an object from its position in that space. This may to some be a radical idea, but it is a common element in real world 'information spaces' such as libraries.

Take, for example, the setting of a notional representation of library books. In many virtual spaces what is important about the data objects is their *position* in the space. This position, is fixed on three axes, and can be used to represent different possibilities. One method used at the moment is to sort these objects in a three-dimensional way according to conceptual similarity. Another method is to indicate people's concerns by mapping what they look at on the Web over the last while and then representing it in its conceptual relation to what other people look at.

By way of a conclusion

By acknowledging the sheer freedom we might conceivably have with regard to representation in a 2 _ D world, we might manage to use this freedom to build the foundations for a 'navigational space' which, in its way would be a kind of mark-up style or convention similar to HTML. The caveats represented by the work above is that we must be extremely careful in how we attempt to represent people and objects in this virtual 'navigational space'. Representation, as pointed out above, can be loaded with political and other problems, which makes the search for an adequate representational style an issue which will be loaded conceptual, political and social issues.

References

1. Heath, C., P. Luff, and A. Sellen. *Reconsidering the Virtual Workplace: Flexible Support for Collaborative Activity*. in *ECSCW '95*. 1995. Stockholm, Sweden: Kluwer Academic.

2. Kasbi, C. and M. de Montmollin, *Activity Without Decision and Responsibility: The Case of Nuclear Power Plants*, in *Distributed Decision Making: Cognitive Models for Cooperative Work*, B.B.J. Rasmussen and J. Leplat, Editors. 1991, John Wiley & Sons: Chichester. p. 275-283.

3. Heath, C.C. and P. Luff, *Collaboration and Control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms*. Computer Supported Cooperative Work, 1992. **1**(1-2): p. 69-94.

4. Heath, C.C., *et al. Unpacking collaboration: the interactional organisation of trading in a City dealing room*. in *ECSCW '93*. 1993. Milan, Italy.

5. Munro, A.J., *An investigation of cooperative work over different media: comparisons to electronic mail*. 1993, Glasgow, Scotland: Unpublished Ph.D. Thesis, University of Strathclyde.

6. Dourish, P. and G. Button. *Technomethodology: Paradoxes and Possibilities*. in *CHI '96 ACM Conference on Human Factors in Computing Systems*. 1996. Vancouver, Canada: ACM, New York.

7. Hughes, J., D. Randall, and D. Schapiro. *Faltering from Ethnography to Design*. in *CSCW 92*. 1992. Toronto, Canada: ACM.

8. Winner, L., *Do artefacts have politics?* Daedelus, 1980(109): p. 121-136.

9. Bowers, J., *The politics of formalism*, in *Contexts of Computer-Mediated Communication*, M. Lea, Editor. 1992, Harvester: London. p. 271.

10. Button, G. and R. Harper, *The Relevance of 'Work-Practice' for Design*. Computer Supported Cooperative Work (CSCW), 1996. **4**(4): p. 263-280.

11. Sproull, L. and S. Kiesler, *Computers, networks, and work*. Scientific American, 1991. **265**(2): p. 84-91.

12. Postman, N., *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*. 1986, London: Hinemann.

13. Mumford, L., *Technics and Civilisation*. 1934, New York: Harcourt, Brace and Company.

14. Benford, S. and L. Fahlén. *A Spatial Model of Interaction in Large Virtual Environments*. in *ECSCW '93*. 1993. Milan, Italy: Kulwer.

15. Bowers, J., J. O'Brien, and J. Pycock. *Practically Accomplishing Immersion: Cooperation in and for Virtual Environments*. in *CSCW '96*. 1996. Cambridge, Massachusetts: ACM.

16. Wallen, L.A. and A.J. Munro. *Multi-Media in Support of Collaborative Work*. in *Presented at meeting of British HCI Group*. 1995. London.

17. Winograd, T. and F. Flores, *Understanding Computers and Cognition: A New Foundation for Design*. 1986, Norwood, New Jersey: Ablex.

18. Garfinkel, H., *Studies in Ethnomethodology*. 1967, Englewood Cliffs, New Jersey: Prentice Hall.

19. Heritage, J.C., *Garfinkel and Ethnomethodology*. 1984, Cambridge: Polity.

20. Schutz, A., *On Multiple Realities*, in *Collected Papers*. 1962, Martinus Nijhoff. p. 207-259.

21. Benford, S., *et al. Shared Spaces: Transportation, Artificiality, and Spatiality*. in *CSCW '96*. 1996. Cambridge, Massachusetts: ACM.

22. Ishii, H., M. Kobayashi, and J. Grudin. *Integration of inter-personal space and shared workspace: ClearBoard design and experiments*. in *CSCW '92*. 1992. Toronto, Canada.

23. Benford, S. and C. Greenhalgh. *Introducing Third Party Objects into the Spatial Model of Interaction*. in *ECSCW '97*. 1997. Lancaster, UK: Kluwer.

Notes

[1] Here I acknowledge Mike Robinson's work on "work as theatre" and particularly a talk given as part of the International Summer School on CSCW, University of Jyväskylä, Finland, August 1997.

[2] This transcript notation is simplified.

[3] The only important parts of the notation for the current discussion is the use of capitals for shouting, the use of double brackets to describe actions, and the use of single brackets to give time, in seconds. Words in brackets are those which are unclear.

[4] The diagram of the flight strip is taken from the paper by Hughes et. al. [7].

Chapter 8

Route Guidance Issues; Verbal versus Map Instructions, and Route Choices

Kristina Höök
SICS

Current work on navigation in electronic worlds is based on the assumption that geographic and electronic worlds are similar enough to make it possible to use results from work on environmental psychology and related areas in the design of electronic information spaces. The present paper is an attempt to analyze the underlying assumptions behind this approach in some detail, as well as an attempt to describe a number of different dimensions on which these spaces can differ. We also discuss how these differences might influence user behavior and design.

Route Guidance Issues; Verbal versus Map Instructions, and Route Choices

Kristina Höök
SICS Box 1263 S-164 28 Kista Sweden
kia@sics.se
http://www.sics.se/humle/

What we want to argue here is that, first, written or spoken descriptions may be much more feasible than maps for many different reasons. Second, the system must have some understanding of how people choose and plan routes in order to both plan the route as well as describe it in a proper manner. Finally, we shall claim that users with different tasks, such as commuting versus driving a taxicab, will have quite different needs in terms of both route choice and description of routes.

Route guidance in the "real" world

Route guidance became a hot issue in the late '80ties in the European research community when the European Union decided to hand out large amounts of funding to come up with new IT tools to be used in the car. Route guidance systems are placed in the car and will give the driver continuos information about where they are and, if the user's destination is provided, they can plan the route and give advice on how to reach the destination. The route can be re-planned if the driver takes a wrong turn. In some systems dynamic information of the traffic situation (accidents, queues, etc.) may influence the route planning. Route guidance systems are now available in rented cars around the world, and may soon become cheap enough to be bought for private cars too.

A problem with the route guidance systems designed is that they are developed by engineers who in general (according to [Borgman 89]) have fairly good spatial abilities and map-reading skills. This may be part in explaining why most of these systems display their information in the form of more or less abstracted maps on small displays in the car.

What we want to argue here is that, first, written or spoken descriptions may be much more feasible than maps for many different reasons. Second, the system must have some understanding of how people choose and plan routes in order to both plan the route as well as describe it in a proper manner. Finally, we shall claim that users with different tasks, such as commuting versus driving a taxicab, will have quite different needs in terms of both route choice and description of routes. Even if these issues relate to navigation in the real world rather than in the abstract world of information, some lessons can be learnt on people's ability to navigate in general. We may also draw some conclusions as to how navigational information may be best presented that may (wholly or partly) transfer to the information world. Among the new areas that are attracting a growing attention are augmented reality, i.e. reality that upon which an information world is superimposed through using various devices, such as wearable computers. Route guidance systems in the car can be considered to belong to these augmented reality applications, and is therefore of great interest to us.

Let us start by discussing the issue of verbal versus map presentations of route information.

Verbal versus Map Descriptions

In studies of navigation in the real world and of route guidance systems, all evidence points in the same direction, namely that verbal instruction rather than maps or pictorial presentations should be used in the car. To some, this may seem counterintuitive, as they may believe that people find their way through cities by using maps. But is it true that people use maps to the extent we believe that they do, and even if they do, is it really the best way to find a destination?

Understanding Maps

Understanding and memorising maps is quite a difficult problem involving many different cognitive procedures. In a series of studies Thorndyke and colleagues, [Thorndyke and Stasz80, Thorndyke and Goldin81], have tried to identify which procedures are used to memorise maps and to depict which of those are used by good map readers as opposed to procedures used by bad map readers. They show that good learners success is largely dependant on their ability to structure an otherwise unstructured task; they also demonstrate a variety of successful techniques for encoding both spatial relationships and verbal labels, and they evaluate what they have learned as opposed to what have not learned so that they concentrate on the right subset of information. Poor learners on the other hand, do not focus on a subset of the information in the map, their procedures for encoding spatial information is poor, ineffective, inappropriate or non-existent, and they do not evaluate what they have learned or not. Instead they often spend time relearning parts of the map that they should in theory already have learnt.

In studies they also discovered that people, regardless of whether they are good or bad learners, learnt more verbal information than spatial. Spatial information associated with verbal labels was also more easily learnt than spatial information without labels.

Scholl and Egeth, [Scholl and Egeth82], have tried to find that underlying capabilities in humans that predict if somebody is a good or bad map reader/learner. They show that predictions of visual-spatial ability and hemispherity are not good predictors of map-reading ability contrary to prediction. Instead it seems like map-reading abilities are primarily dependant upon verbal-analytic ability. They also show that there are possibly different underlying processes for map-reading abilities as opposed to map-learning abilities. (It should be noted that their studies were on topographical maps, which may be different than ordinary maps).

Streeter and Vitello, [Streeter and Vitello86], have investigated how map-reading abilities of different population groups affect their navigational preferences, habits, experiences, abilities and route-selection strategies. They show that self-described good navigators like and use maps and differentially value landmarks, while poor navigators tend not to use maps, prefer verbal instructions and tend to rate all landmarks as equally valuable for route finding.

It was shown that people's self-reports concerning navigational proficiency are to a large degree correlated with objective measures of spatial ability. An interesting difference found between the two groups, was that the low-ability group relied heavily on receiving and giving verbal directions rather than spatial ones.

Among the conclusions drawn from the experiments were that the low-ability group relied on landmarks to such an extent, that they might be said to navigate entirely by landmarks. Their internal map can be described as "a linked list of items, with landmarks as the items". This correlates nicely to studies on the mental representation of spatial information, and how it changes with age and time (see section 2.4.1). The high-ability group on the other hand had both landmarks and spatial cues available to them.

There is also a study on the differences we shall see between knowledge acquired from a map and knowledge acquired from navigation, [Thorndyke and Hayes-Roth82]. From a map people acquire survey knowledge encoding global spatial relations. This knowledge resides in memory as images that can be scanned and measured like a physical map (see further the discussion in section 2.4.1). From navigation people acquire procedural knowledge of the routes connecting diverse locations. People combine mental simulations of travel through the environment and informal algebra to compute spatial judgements. In their study though, Thorndyke and Hayes-Roth show, that with extensive exposure to the city and routes in it the performance superiority of maps over navigation vanishes.

Concerning the question of whether driver locate their destination by using a map, there is an interesting study by Gordon and Wood, [Gordon and Wood70], where it is shown that, in fact, maps are seldom used. Instead drivers adopt the strategy of going to the local area and asking service attendants there for detailed routing information. Six of the 20 drivers in their experiment, started out towards an unfamiliar address without obtaining any advance information at all. Only about 10% used the map (provided by the experimenters) at all.

The studies described above have important implications on our approach to use verbal presentations. Since even good map readers/learners were better on learning and interpreting verbal information than spatial information [Thorndyke and Stasz80] and the verbal-analytic ability was underlying their map-reading skills [Scholl and Egeth82], we get a hint to how difficult it is to transform spatial information to an exact spatial internal representation. We must also consider the group of people that have a low map-reading ability, [Streeter and Vitello85]. Moving maps or even simplified maps in the car will not help them. The conclusions we may draw are:

- *trip planning phase*: If the whole route is presented, it needs to be memorised, so the success of this phase will heavily depend on the drivers ability to memorise the plan for the trip. The studies indicate that verbal information will be more easily memorised and (by some groups of the popul*ation) more easily understood.*
- *route guidance phase*: Since a map needs to be processed for a longer time period in order to be usable to the driver, it will limit the time for processing the surroundings of the automobile.

If spatial information should be used in our system for some reason, it should be associated with a verbal label, in order to make it more easily interpreted and learnt.

Studies of Route Guidance Interfaces

Studies of interface of route guidance systems can be classified and understood in a number of ways. We can look at different ways of presenting information: maps, auditory, simplified maps, etc. We can look at different complexity levels of information: information for the next intersection or several intersections ahead, information content in each message, etc. The way these studies have been conducted will also influence how usable/generalisation they are to us. Some are done in laboratory situations, some compare maps only with combined verbal and pictorial information, etc. The conclusion drawn from these studies will also differ if we want to follow objective results like error rates, eye-movements, etc., or if we want to pay more attention to subjective evaluation. Instead of dividing this chapter into such subsections, we shall only describe a number of studies from all of these aspects, and then try and make some conclusions as to whether verbal information have proved to be better than other presentation forms in studies with route guidance systems.

It should be observed that none of these studies is on resident drivers or on descriptions aimed at resident drivers. Furthermore, all are directed towards determining the best way to present navigational information, not trip planning information.

The most well known studies in this area are Streeter and Vitello's series of studies, [Streeter and Vitello86, Streeter et al.85, Streeter and Vitello85]. Navigational aids where investigated by Streeter et al. in order to determine which were best for navigational information. The aids that were investigated were customised route maps, vocal directions or both. In additions to these groups, there was a control group that was equipped with a state road map plus the address and city of destination. The last group was encouraged to use any help they ordinarily would. The interesting result was that drivers who listened to vocal directions drove to destinations in fewer miles, took less time, and showed about 70% fewer errors than the map drivers. The performance of drivers with route maps and voice directions was between that of the map only and voice only drivers.

Streeter et al. say that one might have predicted a priori that two sources of information would have been better than one. It appears though, as if that was too much information and also that attending to the map for decisions of when and where to turn, produced errors.

The voice direction group got its instructions via a tape recorder that they operated themselves. The instructions were given in four different parts:

1. **Critical direction:** Left or right instruction: Drive for {X.X} miles to {street_name} and turn {left | right}.
2. **Continue instruction:** {street_name 1} changes name to {street_name 2}. At {X.X} miles turn {left | right} onto {street_name 3}.
3. **When to turn instruction:** If a landmark is available, then: {landmark} is on the {left (corner) | right (corner) | straight ahead}.
4. else select the street before on the same side as the next turn. {street_before_name} is the street before {street_name}.
5. **Too far instruction:** If you come to {landmark | major_street} you've gone too far.
6. **Summary instruction:** Remember it's {X.X} miles to your {left|right} turn onto {street_name}.

Observe the negative information given in part three. People tend not to include such negative information when they give route guidance information, but in the study Streeter and Vitello could observe that this negative information reduced the number of errors. The instructions are designed to that the "spacing of practice" principle is applied, meaning that the critical information was repeated, and placed first and last in the message.

In the experiment, no real route guidance systems was used, so, for instance, the voice directions group would not get any help from the tape recorder if they were lost. Despite this fact the vocal directions group performed much better than the other groups.

Parkes, [Parkes90], questions the study made by Streeter et al. concerning the conclusion that verbal instructions are superior. Since they did not test screens with the ability of showing pictures, the comparison between verbal mode and map mode is not fair. Furthermore, Streeter et al. claim that verbal instructions are less visually distracting, but Parkes argues that the processing of verbal instructions like "left" or "straight on" may be more demanding and time consuming than symbols, like for instance arrows, since they need to be interpreted. Since verbal instructions have a transient nature, they might be forgotten when it is time to execute them. The automobile is also a noisy environment that complicates the transmission of verbal instructions.

Parkes claims that the role of memory in the driving situation is a key consideration, when there is a delay between recognition and recall, when stimulus is only available for a limited time or when the information load is high, memorisation is even more important.

He also draws the attention to the problem of understanding maps for a large sub-set of the population and concludes that they should not be used in automobiles. He also points out that 95% of all accidents are due to human factors [Parkes citing Sabey and Stoughton80].

Parkes has together with Coleman, [Parkes and Coleman90], suggested that a different way of investigating how the interface should be designed, is to set-up a similar situation in a laboratory. Here, the authors investigate the effectiveness of different modes of presentation, directional symbols, printed text, and voice simulation. Task completion times, error scores, eye movements, blink rates and subjective measurements were the criteria used for evaluation. Instead of trying a real route guidance situation, they have created a study where the subjects were asked to trace a maze with the help of guidance messages. The messages were either written (LEFT, RIGHT, UP, DOWN), shown as arrows, or spoken. The authors argue that both objective measure of the effectiveness of the modes, as well as subjective measure are necessary when deciding upon which interface is the best. Their results show that voice directions resulted in much faster task completion time and were preferred by the subjects. On the other hand, there were more uncorrected errors with the voice directions, and subjects also felt paced by the auditory commands.

The authors therefore conclude that even if voice directions seems to be the best alternative, they might result in incorrect routes and that people feel paced by the system. A combination of visual and voice presentation seems to be the best.

Labiale, [Labiale89], has conducted two different studies along the same lines as Streeter et al.:s study. Labiale has attempted to investigate the influence of different presentation modalities, visual / auditory / repeated auditory and complexity levels of different in-automobile road information systems. The measurements are subjective preferences and perceptual and cognitive performance.

The levels of information refer to the amount of information, and he has designed four different levels. For example, level 1: "Traffic jam on A40", level 2: "Fog on A31. Speed limited to 60 kph", level 3: "Road works on A62 near Fougerolles. Traffic slowed to 80kph towards Bordeaux" and level 4: "12 km traffic jam on the outskirts of Salbris in the Paris-Province direction. When leaving Orléans follow the yellow itinerary towards Vierzon". (The messages were generated by hand, not by a system).

In the first experiment, he compared three sensory modalities each combined with these four levels. The first modality, which he calls the visual modality, consists of written messages, presented in a segmented format. The second modality, the audio modality, transmitted the message via the loudspeakers only once. The last modality, named the repeated auditory modality, repeated each message 3 s after initial transmission. The memorisation results show that there were no statistically significant differences among the modalities. The number of items remembered, on the other hand, could be closely linked to memorisation, where people in principle are able to remember 7 - 9 items of information, irrespective of medium (which by the way closely connects to short-term memory limits and chunking). People seemed to prefer repeated auditory information, unless the message was very short, in which case people then tended to prefer the visual presentation.

In the second experiment, two variables were investigated. The first concerns what guidance messages are needed as a complement to a displayed map. Either audible or visual guidance was provided. The map itself, was simply a network of white light routes, where the route to be taken was not marked. The second variable investigated was the number of route direction changes, 1 turn or 3 turns. The experiment set-up was such that subjects where only allowed to look at the map, and/or the written message for 30 seconds. It was then taken away.

The results of the second experiment show that overall people remembered map plus written guidance message better than map plus auditory guidance message, especially when the maps become more complex (more than 3 turns). People, on the other hand, seem to prefer the maps plus auditory option. It is shown that the number and duration of visual exploration from the maps reduces with this last option.

In his conclusions, Labiale says that even if oral presentation is best, it needs to be given at the right moment. If not, it may induce mental overload. Repetition of the oral message also seems to be necessary.

In another paper by Labiale, [Labiale90], the second experiment from above is described in more detail. An interesting conclusion in this second paper is, that if the route has been described with words, written or spoken, the structure and itinerary recall will be strongly influenced by that description. For example, the ordering of actions to be taken will be the same as in the description (73% as opposed to 29% when given the map only). The serial property of verbal descriptions versus the parallel property of map descriptions (as perceived by the subjects), is certainly yet another argument for verbal descriptions.

Antin et al., [Antin et al.90], have tried to evaluate the effectiveness and efficiency of navigating with a moving-map display relative to navigating with a conventional paper map and along a memorised route (as a control condition). Results indicated that there were no differences in the quality of routes selected when using either the paper map or the moving map to navigate. However, the moving map significantly drew the driver's gaze away from the driving task. In fact, people spent 33.07% of driving time looking at the moving-map compared with 6.79% looking at the paper map. Since a moving-map system cannot be compared to a route guidance system using a map to display the route, the conclusions we can draw from this study are limited. Even so, it is interesting to notice how much time is spent on looking at the map. According to another study cited by Antin et al., [Hughes and Coles86], 30-50% of the visual attention of the driver may be allocated to things unrelated to the driving task. This implies that no resources would be left to attent to other help and warning system that might be available in tomorrow's car, if a moving-map is chosen.

In [Antin et al.88], yet another attack at the moving map idea can be found. Here, the authors try to find to what extent the driver's spatial ability will influence the process of automobile navigation. They show that in fact performance was not affected by spatial ability, only the total time to derive necessary information from the maps was affected by spatial ability. They also compare paper map presentation with a moving-map system again. They find that since the strategy for using the paper map, is to learn parts of the route while the car is at a stand, it actually promotes procedural learning. The moving-map display in their study did not allow the driver to see the entire area she would travel through. The moving-map display therefore drew a substantially greater proportion of visual attention than did the paper map, and it did not encourage a procedural, more serialised learning.

Alm and Berlin, [Alm and Berlin91], have studied the amount of verbal information that produces the best results in terms of subjective evaluation. They set up three levels of amount of information, where the first only gave information about the next intersection coming up, the second and third level included instructions about what to do in the two respective three next intersections. The instructions were tested with drivers unfamiliar with the area, and showed that subjects preferred information about the next two or three intersections, but that the information should depend upon driving times between intersections. The rate of repetition demands, showed that information about the next three intersections was slightly too much. Some of the subjects also asked for more information in situations where they should simply go straight ahead: "go straight ahead until..." whenever the driving times until the next intersection were above the median. The authors conclude that landmarks could be used to construct these straight-ahead messages.

Alm and his colleagues, [Alm et al.91], have studied the use and effects of landmarks further. Basically, they found that there was a positive effect of using landmarks in the descriptions. It made the drivers more certain that they were at the right spot. From the study they also conclude that a combination of spoken and visual information is probably the best. The argument being that the driver should be allowed to confirm what she thought she heard whenever she desires to.

Parkes and Martell, [Parkes and Martell90] have studied drivers using a PC-based route-planning product. They found that about 70% preferred the text-based route guidance system, as opposed to their normal method for route finding. Their normal way of route finding typically include finding their way through taking notes from a map and confirming their progress by reference to landmarks. It should be observed though, that their PC-based route planner did not include landmarks, and no real-time traffic information.

In a paper by [Kramer and Reichart89], something called the "commando effect" is discussed. The command effect comes into play when the driver perceives the instructions from the system as commands. Instead of following their own common sense about when and how fast to act, they try and fulfil the commands as quickly as possible. Something of that effect has been informally reported about the LISB system, where especially elderly people felt paced by the system instructions.

An interesting completely implemented verbal interface to a route guidance system can be found in the Back Seat Driver [Davis89]. The system is foremost aimed at tourists, and includes both instructions of turns with landmark descriptions, etc., as well as advice upon which lane to take and which speed to hold. Unfortunately, we have not found any formal studies on whether it has been a success or not, there is only some anectodical evidence in Davis Ph.D. thesis. The Back Seat Driver will be described on more detail in section 4.4, as a comparison to the system we have implemented at SICS.

Unfortunately, there is not one single study that tells us whether maps, or verbal (either audible or written) messages are to be preferred. Possibly, we cannot design such a study, because there are so many other factors to take into account. When Streeter concludes that verbal instructions are best, the competing map condition was not a moving map, or a map with the route highlighted, it was simply a paper map. In Parkes and Coleman's study, we are faced with a laboratory situation, which cannot be compared, to the real-life situation in a car. Labiales two studies have compared some more combinations of maps and symbols, but unfortunately, he never actually compared maps with symbols. In the first experiment he only compared various symbolic messages with one another. In the second study, maps were used in both conditions, with written versus audible messages. Antin et al. studied moving-maps versus paper-maps.

But even if we cannot find a simple answer to the question of which form of presentation is the best, all these studies together help us to form a picture of the situation and the factors that influence it.

It seems fair to take as a working hypothesis that verbal presentations are a good starting point, both from a map-reading/learning point of view (section 2.3.1) and as a conclusion of the studies described in this section. Repetition also seems to be an important factor. Repetition can be provided if the verbal route descriptions are written on a screen in the car, as well as being presented audible. The difference between trip planning and route guidance is important, and closely connected to how much information is contained in the instructions and how many actions ahead are described. Not only how much information is contained in the verbal descriptions, is important, but also what information; landmarks, advice on lanes and speed, etc.

Human and Computer Route Choices

With a perfect route guidance system, we could claim that the driver should simply sit back and follow the directions given, and she will get to the destination on the quickest, shortest, or "most scenic" route without having to plan or understand the route herself. Unfortunately, the shortest or quickest or even the most scenic route might be horrible to drive. It might contain many left-turns, it might go on small, narrow roads, etc. The driver might also be unwilling to simply follow the direction, instead she will abandon a route suggested by the system if it diverts too much from her own choice. It is therefore necessary to study what kind of criteria humans adopt when choosing routes.

Mental Maps and Representation of Spatial Information

Underlying human route choices are human capabilities of representing and processing spatial information. In order to decide upon the issue of 'multiple intelligences' (which [Gardner87] argues to be the conclusion that can be drawn from the fact that completely different kinds of representations and processes are used for imagery and other kinds of thinking), lots of studies have been conducted where different aspects of images have been examined. Since the approach taken in here is that it is a good idea for the resident driver to get an overview of the route before take-off, what we are asking for is really that the driver forms a mental image in her mind of the route to be travelled. The time needed for doing this, and the relative importance of the information units we give her, can be investigated through studying the experiments on imagery, mental maps, etc.

Research into mental maps, has been very much concerned with what kind of knowledge representation humans use when encoding spatial information. [Kosslyn et al.79], believes in the quasi-pictorial representation, usually referred to as 'the-map-in-the-head' metaphor. Others like [Pylyshyn81] claims that the underlying representation is always propositional.

Kosslyn has argued in several books, [Kosslyn80], and papers that humans are capable of a certain kind of activity that he calls imagery. This capability is according to Kosslyn, separated from other kinds of thinking that humans are able to perform in several important ways. He means that the representation of these objects differs from representations of for example, verbal information. This is challenging "the language of thought" research which argues that there is only one representation for all kinds of information in the human mind, and that that representation is some kind of propositional language. Since Kosslyn argues that the datastructure is different, he argues that it follows that the kinds of procedures that are applicable on this datastructure are different from other procedures. For example, we can apply transformation, visualisation, etc., on them.

In [Baird et al.82], we find an attempt to apply a mathematical theory of distances to results of studies of human cognitive processes. If the internal human representation of the physical world is a two-dimensional cognitive representation, a mental map, some of the mathematical properties they put up should hold, but they do not. (Examples of mathematical properties are triangle inequality, that the distance from x to y equals the distance from y to x, etc.). The authors then put forward the theory that our internal representation is not a perfect Euclidean mapping of the external world. Instead we first transform the impressions we get, that representation is then transformed to yet another internal form, and we then transform this internal form again when asked to judge distances, or do spatial mapping, see figure 4. Baird et al. draws the conclusion that experiments on humans internal cognitive representation of spatial information, will basically be very hard to do, since what we can measure is only the two surface representations, not the underlying, internal memory representation.

This would have serious implications on what we are trying to do in this section of this thesis. If we allow the limitations of what we believe to be the internal representation of spatial information to influence our design of the route guidance interface, and it turns out that what we have studied and used as a basis for that is only the surface representations, our assumptions might be wrong. On the other hand, the internal representation of spatial information only has a secondary influence on our design. It will only put up some limitations to what we can expect humans to be able to process, and process quickly.

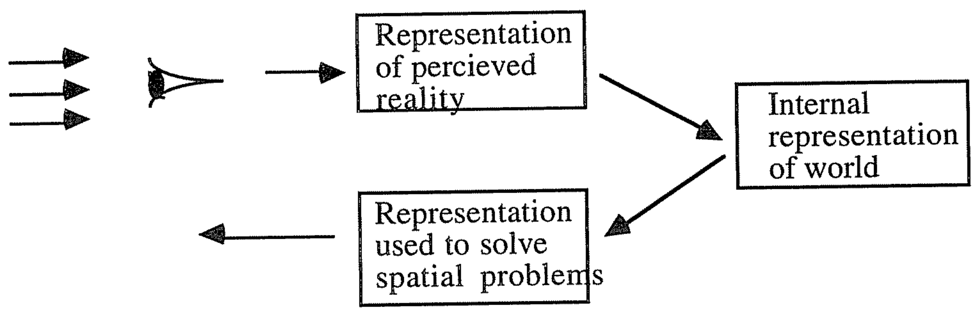


Figure 4. Representations of spatial knowledge.

Another attack by [Kuipers82] on Kosslyn's map-in-the-head metaphor is maybe more serious than the mathematical evidence for deficiencies of the representation. Kuipers is interested in the stages that humans go through when forming the mental map. He says that we should go back to a notion of two different representations. One more fundamental representation that is learnt easily by children and another that you need to train yourself to be able to make.

The first representation simply involves the ordering of objects you pass on, for example, a path, the connectivity and containment. Whenever passing an object this gives the clue to what we expect next and what action to perform right now. So that we remember when seeing the hedge, that it is this corner where we should turn left. We do not have a mental representation of the entire route in our head that we can reconstruct in our mind, but we rather have a list of things that comes alive, part by part when passing objects in the list. In this representation we are unable to measure time and distance or to relate objects not visible for the moment to one another, directions and distances between them etc. This representation might explain why people are sometimes unable to go from B to A even if they know how to get from A to B.

The second presentation is one where metrical relations, distance and direction, are available. This representation is not learnt until later, and it is helped by viewing a map, and through extensive experience of the area, see section 2.3.1, [Thorndyke and Hayes-Roth82].

Kuipers claims that if we really have a map representation in our head, we would be able to relate any object in that map to any other object in the map. Since we frequently are unable to relate one region to another, but might have a perfect orientation within the regions, we are sometimes unable to judge correctly how an object in one region relates to another object in another region. Kuipers proposes that modifying the map-in-the-head-metaphor into an Atlas-metaphor where one region is represented as one page in the Atlas, and two regions are related only at a few point where they connect, would be a better metaphor. So instead of assuming the existence of an isomorphic function that converts all points in reality into points in the head, and that possibly gets some of these points wrong, we can predict that an entire region might be somewhat rotated compared to another region. This would explain why all the points in one region are distorted in comparison to the other region.

In [Gärling et al.82], the development of mental maps is also investigated. They show that, contrary to the prevailing hypothesis, people seem to learn paths between landmarks, before they learn the relative locations of landmarks. More important to us, is that they show that the relative directions of landmarks are quite easily learnt, but that the relative distance to landmarks is much harder to learn.

In another study by [Sadalla and Magel80] subjects estimated distances along a route containing varying numbers of right-angle turns. Based on memory of the route, estimates of distance increased with the number of turns. Subjects seemed to simply sum up the number of segments travelled between turns, rather than estimating the length of those segments and adding them up. It is also known that people's estimates of distances from x to y sometimes are different from their estimates of distance from y to x. Similarly they violate triangle inequalities.

So, when describing routes, we shall try to avoid using quantitative measures like kilometres, number of blocks, etc, since, in general people are very bad at estimating distances.

In general, the exact way that spatial information about a city is represented in the brain will not affect the design of a route guidance interface. What will affect the design, though, is what limitations, the representation will have on human performance when solving spatial problems. We see that human are not good at estimating distance, or relative location. We also know that humans in many cases choose suboptimal routes.

The stages human go through in order to learn spatial information, will also have an impact on the route guidance interface. Since "connectivity" is the first way of learning, an obvious conclusion is that some instructions cannot be given until the driver sees the intersection or place where the action should take place. It might also provide a basis for making intelligent guesses about aspects of route description that might be misunderstood by the driver.

Human Route Choice

There are some studies on how humans choose routes. These investigations into trip planning are mainly concerned with how people plan routes in terms of the criteria they find important and how they search the map, or their memory to construct the plan. Investigations into how to design trip planning systems from a cognitive point of view was first investigated by [Elliott and Lesk82].

Elliott and Lesk implemented a system, in which they could, given a starting point and a destination, plan a route between the two points. They have implemented several algorithms in order to test their feasibility; breadth-first/depth-first search, pre-storing important routes, divide-and-conquer, and keeping a hierarchy of maps with progressively fewer streets. Feasibility was tested both in terms of fastest algorithms, but also in terms of the route being understandable to humans, which is what we are interested in here. People's strategies seemed to be first finding any main roads, and then apply divide and conquer as well as depth-first search.

The experiment set-up consisted of eight subjects. They were asked to sit down with a map over an unfamiliar area and decide upon routes between two points. In general people first searched for an important road going in the right direction. Then they did depth-first search, single ended, going to and along important roads. A first-hit strategy was used, they did not look for a second route. Interesting enough, people tended to depend upon colour and presentation, not labels, to find main streets.

Elliott and Lesk conclude that a hierarchical search probably will be the best computer algorithm for finding routes that are both short and acceptable to humans. It was shown that the theoretically best algorithms produced routes that took too many turns, and they were also unfeasible since they were designed to handle the worst-case, rather than the normal case.

In another study [Streeter and Vitello86] found that a good predictor of whether people would chose a particular road was whether the sum A + B + C (were A equals the straight-line distance from the start to the road, B equals the distance traversed on the big road, and C equals the straight-line distance from the departure point on the road to the destination) did not exceed the straight-line distance between start and destination by more than about 20%.

They also showed that residents used primarily the local road system, whereas experienced subjects (living in the area) used secondary roads and major roads, and tourists used mostly major roads. Humans seemed to follow a hierarchy of the road net, where they tried to find a road at one of the levels in a global way, i.e. looking at all the roads at that level. All the algorithms on the other hand, used a local way of searching a particular area.

Humans tended to switch to a lower level in the hierarchy when the difference between A + B + C (see above) and a straight-line distance exceeded 20%. For tourists, a difference of 30% is acceptable.

In a follow-up study Streeter and Vitello investigate how a system can find routes that people unfamiliar with an area find usable, [Streeter and Vitello85]. They have studied experienced rental automobile employees, who characteristically give customers directions. They compared routes generated by their own heuristics to the automobile rental employees' choices.

There are two interesting properties of Streeter and Vitello's heuristics. The first is that it uses a birds-eye view when planning the route. They claim that it is reasonable that humans look further than to the next intersection when they plan a route, and that a more global view is necessary. In fact, only in 5-15% of the cases other algorithms succeed in matching humans choices, while their algorithm accounted for 52% of tourist routes, 33% of experienced and 17% of resident routes.

The other characteristic of their algorithm is that it uses a three-level hierarchy of the city. One level is the local road system, the second is the secondary roads and major roads, and the last level are major roads. People seem to try and solve problems at the highest level first, and then going down in the hierarchy (see above). As expertise evolves, this structure is flattened, so that an resident searches through the entire network, instead of just one level.

In this study, Streeter and Vitello again find that people tend to stick to as high a level in the hierarchy as possible, even when this leads to a considerable loss in terms of distance.

A question to be asked in this context is whether the routes generated by residents and directed towards tourists, are necessarily the best just because humans choose them. Streeter and Vitello argue that since these routes are the most easily described, they must also be the best for tourists. If one imagines a perfect route-guidance system that is always able to give perfect and unambiguous information, it might not be necessary to chose routes that are easily described. Instead, we might imagine selecting routes that are 'better' in the sense of being shorter in time or distance.

Both [Chase83] and [Pailhous70] have made investigations into how taxi drivers learn and make use of their knowledge of large-scale environments.

Pailhous divided the streets of Paris into a 2-tiered hierarchy: a base and a secondary network. The base network was defined as those streets that were highlighted on the Paris map, about 10% of the streets. The secondary network became the rest. Pailhous studied experienced and novice taxi drivers and found that both groups used the base network. When presented with a detour problem, Pailhous found that half of the expert taxi drivers would chose a route through the secondary network to get around the barrier in an optimum way. Novice taxi drivers on the other hand, would select a longer base network route to get around the barrier. Pailhous concludes that the basic strategy of taxi drivers is to get to the base network as quickly as possible, and then stay on it for as long a possible.

Chase has studied Pittsburgh taxi drivers, and found results that to some extent contradicts Pailhous. The expert taxi drivers in his study, did use the secondary network whenever possible, instead of sticking to the base network. Chase gives two possible explanations to this: a) there is no hierarchical division of the street system (with the exception of the Parkway in Pittsburgh) and streets vary on a continuum of familiarity, or b) the preferred street system, or base network, expands with expertise.

Chase furthermore argues that taxi drivers do not navigate by means of a map in the head, instead the underlying representation of the city is hierarchically organised so that locations are nested within large regions and neighbourhoods, neighbourhoods are nested within large regions and larger regions are located with respect to more global features. Chase suggests that the hierarchical storage of places and streets also offer economy of storage, and that it is an integrational part of planning a route. To get from one location in one neighbourhood to another location in a different neighbourhood it is suggested that the driver first finds a route that connects the two neighbourhoods, and then the rest of the route is either subsequently generated or it is filled in as the driver goes along. The driver can continue to follow the "global" plan until cues from the environment are encountered that trigger specific routes at choice points along a route.

As we shall see in the next section, a failure to take human route choice into account in a route guidance system might prove to be the crucial point of acceptance of that system. A route guidance system should take the hierarchical view of the city into account, either for the actual planning of the route, or for the presentation of routes that do not follow the normal pattern of hierarchy. Also, the A+ B+ C theory helps us understand when a route chosen by the system might be rejected by the driver.

In general, the human route choices need to be considered either by the route planner, which may or may not adapt to them, and/or the route presentation that can emphasise parts of the route or explain why a certain route is better. The route presentation can only do those kinds of explanations if it understands what human route choice means. This is further investigated in [Lindevall and Höök91].

Evidence from Evaluation of Existing Route Guidance Systems

Above we have outlined some of the criteria humans find important when choosing routes, and we have also tried to find what underlying capabilities that restricts human route choices. Even so, we might ask whether it is really necessary to take these into account when designing the route guidance system? A study, [Bonsall and Joint91], on the LISB system and a laboratory system called IGOR, shows that it is actually necessary to take human criteria of route choice into account. The LISB system does not take human preferences into account, and use of the system diminishes rapidly after a time of use.

First, a short review of the LISB system: the user starts by keying the location of their intended destination into an in-vehicle unit. A small screen then displays the crowfly direction and distance to the destination (known as "autonomous mode"). When the vehicle passes a roadside beacon, the display changes to "full guidance" mode. From this point on, LISB's calculation of the minimum time route to the destination is made known to the driver via symbols on the display and audible messages instructing the driver to make the requisite turning at each intersection. Towards the end of the journey, the system reverts to autonomous mode. (The autonomous mode is due to the lack of enough beacons, or rather possibility to handle more than a limited number of beacons. There are therefore only beacons in the largest intersections).

In the study of the LISB system, they found out that the use of the system seriously diminishes after a period of use. The main obstacle for using the system seems to be the way the destination is given to the system. The study also shows that even in cases where people actually use the system, about 30% will still not follow the instructions. The reasons for not following the instructions were according to the subjects:

- they thought that the system was sending them in the wrong compass direction,
- the advice was given too late,
- the system suggested leaving a route which was normally good and which had no obvious problems the day in question,
- it suggested they use a route that was normally very congested,
- the system was apparently malfunctioning,
- it sent them in a direction contrary to the road signs.

The authors conclude that all of these demonstrate that, unless guidance is backed up by other information available to the user, either from her experience or from her direct observation of prevailing circumstances, she may well decide not to follow it.

The authors further conclude that drivers are unlikely to request guidance if they find the effort of doing so out of proportion to their perceptions of the benefits to be gained, and that they are very likely to ignore or reject guidance advice if they do not find it credible. This fits with the theory of human effort described in section 2.2. Factors affecting its credibility include:

- the extent to which it is corroborated by, or in conflict with, local evidence about the alternative,
- mismatch between the system's route choice criteria and those of the driver,
- the driver's familiarity with the local network,
- the quality of advice previously, and particularly very recently, received from the system by the driver (objective measures of previous quality seem a reasonable proxy for what is no doubt a much more complex phenomenon including factors such as the frequency of obvious malfunctions, etc),
- the driver predisposition to accept/reject advice.

The last point above comes from the second study, where the interviewers were able to classify the subjects into four groups:

(i) dissenters (about 10% of participants) who object to guidance on ground that it is an intrusion on personal freedom;

(ii) the undecided (about 40% of participants) who are torn between a desire to be guided and a cynical mistrust of the system;

(iii) the prudent conformists (about 40% of the participants) who will follow advice if, and only if, it is logical;

(iv) the trusting (about 10% of participants) who are unsure of their own skills and, lacking experience, are happy to put themselves in the system's hands.

Apart from pointing out the importance of route choice, and of how that route choice is presented to the driver, this study also points forward to the issue dealt with in section 2.5, namely user modelling. We need to know how familiar the driver is with the local road network, and the drivers' motivation for using the system in the first place.

User Modelling

A number of authors, [Parkes90, Alm89, Michon85, Michon89] have pointed out how important it is that human factors, especially from a cognitive science point of view, are taken into account already at the design phase of a route planning and navigational system. In what way should we allow human factors to influence design?

[Parkes90] argues that a careful analysis must be undertaken to find out what information needs different drivers have. Different drivers' needs, capabilities and expectations should dictate different mediums and styles of information transfer. He points out that driver information needs are not homogeneous: "Different drivers have varying requirements from a system, e.g. a courier requires the quickest route, whereas a holiday maker may want to follow the most scenic roads. An expert user of a system may require much less detail than a novice user. The needs of an individual driver may well vary over time. For example, whilst travelling on a motorway the driver requires little information, but when entering a city centre, or nearing the final destination, much more detail is required".

[Alm89] has taken available literature on route guidance presentation, people's representation of environmental information, attitudes towards route guidance systems, etc., to serve as a basis for suggesting what kind of functionality, from a human factors point of view, a perfect route guidance system would have. Such a system, would (among other things)

- know about the drivers personal preferences and adapt to them,
- be silent and only provide information when necessary or when the driver asks for it,
- recognise dangerous situations and be silent until the danger is over in order not to increase the drivers mental load,
- be able to identify the special needs of, for example, the commuter and learn her normal route,
- be able to learn which areas the driver is knowledgeable of and which she can be considered to be a tourist on and therefore needs more help with,
- etc.

Neither Alm nor Parkes provides any solution to how these systems should be constructed to enable them to fulfil the criteria they put up. Michon [Michon85] makes a proposition that would get us one step in the right direction. He claims that it is necessary to have an explicit cognitive user model. Michon, who comes from another field of interest; modelling the driving task in order to increase safety, points out how important it is that this model contains concepts like driver intention, goal, etc. In his field, any modelling of humans have been behaviouristic. The driver has been looked upon as somebody from which the system gets "feedback" to react upon. Michon argues that the human actor should instead be regarded as an intelligent, although maybe not infallible problem solver, with whom you interact and cooperate rather than get feedback from. If we want to make a good adaptive system we need to go from adaptive control of behaviour to adaptive control of thought.

In order to do so, Michon, [Michon85 and Michon89], proposes that the driver model should be made explicit and that it should contain concepts like intention, goal, etc. It should also be formulated in such a way that a hierarchical representation is possible, both because humans exhibit such a hierarchical behaviour, but also so that the system can 'learn'. What Michon's means by learning is a follow-up to Anderson's ideas, [Anderson80], about procedural and declarative knowledge, where you allow one bunch of actions to be composed into one single procedural rule that exists on a higher hierarchical level. This would mimic the human behaviour, where we, for instance, first learn to ride a bicycle be learning each step, but once learnt, we are unable to depict what those different steps were.

We can draw the conclusion from these three writers, that we need to incorporate some kind of driver model into our system, and that it preferably should not be just a behaviourist model with stimulus-response pairs as only means of predicting and adapting to human behaviour. The only way we can possibly make a system that can adapt to a user, and learn from a user is if we have an explicit model in the system. In our view, such a model must be allowed to influence more than simply the interface to a route guidance system, it must also affect how the planning is done, how the map database is organised, etc.

In [Wærn et al.90] user modelling is divided into three different levels. The first is to incorporate some general knowledge of human cognitive characteristics, in the sense of the system knowing what a human can and cannot do, not implying that the system itself should be able to simulate human capabilities.

A second level allows for some differentiation between users: using stereotypes. Users can be different on the scales of being expert-novice, being in favour-against, previous domain knowledge, learning style, etc.

The third level is to assess the characteristics of each individual user during the interaction with her. One way is to assume that the user does not have the relevant knowledge when she answers incorrectly and that she has it when she answers correctly.

Apart from these levels, there are two methods, [Nwana91], for identifying faulty answers. One is to compare the user's answers to some bug catalogue of previously known bugs; the error based approach. The other is to have a normative model of how the user should behave, and detect any deviations from that.

Choosing between these levels for our application, we can observe that [Wærn et al90] point out that an individual user model is really only possible when the interaction with the user is quite rich, preferably in natural language. Only then is it possible to diagnose each recipient individually. Wærn et al. refer to [Moore and Swartout88] who propose that only a dialogue where the user has a chance to ask for more information and to offer her own interpretations can provide an appropriate basis for providing a good explanation.

In our application, the interaction with the driver is very limited; we can only observe her reactions, i.e. whether she is following or not following our proposed route-plan. Building an individual user model from that is probably impossible. Using a normative model of the driver requires that we have a normative model of the driver that we do not have (at this time).

The two first levels on the other hand, seem almost necessary for this application. The first level will be realised through designing the interface to the driver in such a way that it takes mental load, human perceptual capabilities, short term memory limitations, etc., into account. The second level can be realised through having different functionality available that are adapted for different user groups needs, like elderly drivers, tourists and residents, commuters, etc.

Let us look at two studies on user groups with special needs, the commuter group, and the elderly drivers group.

Commuters

In [Gray et al.], we find an attempt to define the functional requirements for an advanced driver information system (ADIS), through an extensive study and interview with commuters. The study is quite interesting since it is one of the few investigating commuters' requirements from a computerised traffic information system.

The study is divided into three parts, and conducted in Seattle, Washington, where there are huge congestion problems. The first part of the study was a questionnaire with 3 972 respondents. The purpose of the questionnaire was to determine what kind of groups of commuters could be found. Four groups were distinguished that differ in the degree to which they can change route, or time to travel.

From the 3 972 subjects, 96 were chosen to represent each of the four groups, and were interviewed. The interview shows that commuters are familiar with alternative routes, but they rarely use them, and whenever they do use them, they experience more stress than when they travel on their primary route. It also seems important to see the actual congestion before changing route, and commuters do not seem to believe in the available traffic information, which is information broadcasted on radio. From this questionnaire the authors also conclude that traffic information needs to be specific in term of time and geography.

From these results, together with existing screen designs and human factors guidelines, the team then designed five different interface screens. They differed both in layout, and in information content. These five screens were printed out on paper and sent to the 96 subjects. They were asked to evaluate these screens, and rate them. It was shown that a screen that contained a picture of heavy traffic with time estimates, was the most preferred. When the subjects were asked to rate the information content irrespective of which screen the information was given in, they showed that (on a scale from 1 to 5 where 1 is the most preferred) time estimate were the most preferred (1.97), followed by text messages (2.32). The text messages contained information like "Blocking Accident I-5 & Green Lake – exit 85th to Aurora". Pictures of the actual traffic were ranked high (2.86), while maps (3.17) and bar graphs (4.51) were not ranked as high. The authors draw the conclusion that time estimates of different route alternatives, in the form of numbers rather than pictures, map coding, or descriptive text messages, may be the most effective format for this information.

An interesting change of mind occurred. When people were asked whether they would use a TV-based information service that provided screens similar to the ones showed in the last study, 81% answered yes, while only 25% answered yes in the initial study where they had not been presented with a possible interface to such a service.

Elderly Drivers

In a comparison between the LISB and the TRAVELPILOT systems [Fairclough et al.91], a group of users with special needs was singled out, namely the elderly drivers.

The LISB system has been mentioned above. The TRAVELPILOT system is not a route guidance system, but rather a Route Navigation system. It does not provide any route, but only the information needed to decide upon a route, through a moving map and a positioning device that points out in the map where the car is right now.

The study shows, not surprisingly, that LISB is subjectively preferred, and also that more objective measures like eye movements and other measures of stress singles out the LISB system as a better system in terms of not increasing the human workload too much. Interesting with the study is that elderly people had very specific problems with both systems. Firstly, they had difficulty viewing the screen with both systems. The elderly eye is less able to accommodate for close objects, and using reading glasses is uncomfortable when driving. Some of the information on both types of screen, elderly people would misinterpret or completely miss. Secondly, the elderly subjects also felt paced by the LISB system. When the voice direction command comes, they feel obliged to immediately follow it. Some complained that the command came too late, too close to the point where they were supposed to turn or change lane.

Summary

If we come back to the three hypotheses posed in the beginning of this chapter, we see that we have found some evidence in literature for all of them. At the same time, we have discovered a number of factors that will determine the success of each of them.

The first hypothesis, that verbal direction (written or spoken) is feasible, we have shown to be more or less valid. There are a number of issues to take into account though, like:

- is the presentation repeated somehow?
- how much information is contained in each of the verbal instructions?
- will (especially elderly) drivers feel paced by the system?
- even if human route choices will be generally enhanced by the system, subjects might not "like" the system, so subjective evaluation is important,
- how much attention will the presentation draw from the actual driving?
- etc.

There are several factors that limit how general the results of these studies are. For example: they do not deal with different driver groups, but only tourists; they have not differentiated between trip planning and route guidance and the influence trip planning information can have on the success of the route guidance interface; there has not been a "fair" study comparing a map interface with a verbal interface; etc. Still, they all point in the same direction, namely that spoken instructions are probably of better use to drivers.

We have also found some evidence for the second hypothesis, that human route choice is important to consider. Especially, after considering the study by Bonsall and Joint, it has become obvious that more careful though has to be given to how the route guidance system chooses and describes routes to the driver.

Finally, the last hypothesis from the beginning of this chapter, that an explicit user model is needed, has also turned out to be true, although we have not found any good solutions to what such a model would contain and how it would be realised in the system. Some guidance can be found in [Wærn et al.90], where we find that for this particular problem, it seems reasonable to assume that a stereotypical user model, will both be sufficient, and furthermore the only possibility if we assume a very restricted dialogue with the driver. Interesting alternatives are available if we consider clustering of many driver behaviours, as is done in some machine-learning methods [Maes, 1997].

Implications for Information World Navigation

Even if the discussion in this paper was on navigational aids for the car, we can draw some conclusions on navigation tools for information worlds. This is especially true as route guidance is an excellent example of an augmented reality where the abstracted route guidance information is superimposed onto the real world of navigating in the car.

First, just inserting an abstract map in the interface will be of no use to a large group of the population. There seems to be some evidence in the studies described that:

- verbal descriptions work better since those are easier to remember, processed by other parts of the brain (not requiring so much of our spatial ability and abstract reasoning skills)
- social interactions happens fairly often when people are lost – the ask for advice in gas stations etc. thus supporting the claims made in the PERSONA project that social aspects of navigation should be considered.

In the discussion on route choice as done by humans and machines, it became clear that human route choice is done differently from how machines do it. If the way it is done by machines does not converge convincingly with how people do it, we run the risk of alienating users. Is the same true for automatic aids for information space navigation? We need to get better characterisation of how people perceive navigation in information space in order to determine what are the "normal" ways of doing it. As an information space does not necessarily have all the properties that reality has, as for example, Euclidean distances, measureable distances, landmarks, relative positions of landmarks and places, etc., we cannot expect to see exactly the same behaviours. Maybe users of information spaces are more easily convinced that the advice given is useful since the tools that are available to them are so poor that they cannot really judge the whether a suggested "route" through the space is optimal/conforming with their preferences/desired. For some applications as e.g. a word processor, on the other hand, users might very well be aware of the best way to "go to" a particular "state"/place in the application since the have used the system often.

The discussions on different reasons for using a route guidance system is also transferable to information worlds since it questions the idea that all drivers will have *one* given destination and *one* optimal route that will take us there. The commuters with one given destination and one optimal route that only in extreme cases should be altered, can be compared to the knowledgeable information seeker going to a particular place in the space every day. S/he only needs to be informed of unexpected changes in the space (a server is down so the connection does not work, the disk is full, etc.). The tourist traveller, is a good illustration of the fact that the route to the destination might be just as interesting as the destination itself. A scenic route might in fact change the tourist decision on destination. The same thing happens with the user who is browsing with a more or less precise goal/destination in mind. The more "scenic" the route is to the destination, the better? The taxi driver is perhaps what comes closest to the more traditional view of information seekers. The taxi driver knows where s/he is heading and wants the quickest route to destination. The whole scale of different tasks and needs between these most extreme cases exists of course.

References

[Alm89] "Show me the way to go home - or drive me home old chip!", H. Alm, Available from Swedish Road and Traffic Research Institute (VTI), Sweden, 1989.

[Alm and Berlin91] "What is the optimal amount of information from a verbally based navigation system?", H. Alm and M. Berlin, A DRIVE Report, project V1037 GIDS, Available from Swedish Road and Traffic Research Institute (VTI), 1991.

[Alm et al.91] "The Effects of Landmark Presentation on Driver Performance and Uncertainty in a Navigation Task - a Field Study", H. Alm, L. Nilsson, S. Järmark, J. Savelid and U. Hennings, Available from Swedish Road and Traffic Research Institute (VTI), 1991.

[Anderson80] "Cognitive Psychology", J.R. Anderson, 2:d edition, W.H. Freeman and Comp., New York, 1980

[Antin et al.88] "The Effects of Spatial Ability on Automobile Navigation", J.F. Antin, T.A. Dingus, M.C. Hulse and W.W. Wierwille, Trends in Ergonomics/Human Factors V, F. Aghazadeh (ed.), pp. 241-248, Elsevier Science Publishers North-Holland, 1988.

[Antin et al.90] "An Evaluation of the Effectiveness and Efficiency of an Automobile Moving-Map Navigational Display", J.F. Antin, T.A. Dingus, M.C. Hulse, and W.W. Wierwille, International Journal of Man-Machine Studies, 33, pp. 581-594, 1990.

[Baird et al.82] "Impossible Cognitive Spaces", J.C. Baird, M. Wagner, E. Noma, Geographical Analysis, 14, 204-216, 1982.

[Bonsall and Joint91] "Evidence on Drivers' Reaction to In-Vehicle Route Guidance Advice", P.W. Bonsall and M. Joint, 24th ISATA Conf., pp. 391-401, Italy, 1991.

[Borgman 89] "All Users of Information Retrieval Systems are not Created Equal: An exploration into individual differences", Borgman, Christine L. *Information Processing & Management*, Vol. 25, No. 3, pp. 237-251, 1989.

[Breuker90] "EUROHELP: Developing Intelligent Help Systems", J. Breuker, EC, Kopenhagen, Amsterdam, Manchester, Leeds, available from univ. of Amsterdam, 1990.

[Brown and Lindevall91] "Directed Hierarchical Route Planning with Dynamic Information", C. Brown and P. Lindevall, in "Final Report on Interactive Route Guidance, 1988-1991", SICS Technical Report T91:18, also in proceedings of 24th ISATA Conf. in Italy, 1991.

[Brown and Wærn91] "A Conceptual Architecture for Dialogue Management", C. Brown and A. Wærn, in "Final Report on Interactive Route Guidance, 1988-1991", SICS Technical Report T91:18, also presented at "Nässlingen workshop 1990", 1991.

[Brown et al.91] "A System Architecture for an Interactive Route Guidance System", C. Brown, K. Höök, P. Lindevall and A. Wærn, in "Final Report on Interactive Route Guidance, 1988-1991", SICS Technical Report T91:18, 1991.

[Brown et al.90] "Using a hierarchical Map Database for Route Planning", C. Brown, A. Wærn and P. Lindevall, in "Final Report on Interactive Route Guidance, 1988-1991", SICS Technical Report T91:18, 1990.

[Chase83] "Spatial Representations of Taxi Drivers", W.G. Chase, In D. R. Rogers and J. A. Sloboda (Eds.), Acquisition of Symbolic Skills, pp. 391-405, New York, Plenum, 1983.

[Davis89] "Back Seat Driver: Voice Assisted Automobile Navigation", J.R. Davis, Ph.D. thesis at Media Arts and Science section at M.I.T., USA, 1989.

[Davis and Trobaugh87] "Direction Assistant", J.R. Davis and T.F. Trobaugh, Speech Research Group Technical Memo 1, The Media Laboratory, MIT, 1987.

[Davis and Schmandt89] "The Back Seat Driver: Real Time Spoken Driving Instructions", J.R. Davis and C.M. Schmandt, Proc. IEEE Vehicle and Navigation and Information Systems Conf., Toronto, Canada, 1989.

[Dornic77] "Mental Load, Effort, and Individual Differences", S. Dornic, Report from the Dept. of Psychology, University of Stockholm, 1977.

[Elliott and Lesk82] "Route Finding in Street Maps by Computers and People", R.J. Elliott and M. E. Lesk, AAAI-82 Conf., pp. 258-261, Los Altos, CA:American Association for Artificial Intelligence, 1982.

[Fairclough et al.91] "A Comparison of Route Navigation and Route Guidance Systems in an Urban Environment", S.H. Fairclough, M.C. Ashby, K. Lorenz and A.M. Parkes, Proc. of the 24th ISATA conf., pp. 659-666, Italy, 1991.

[Gardner87] "The Mind's New Science: A History of the Cognitive Revolution", Howard Gardner, Basic Books Inc., New York, 1987.

[Gray et al90] "The Design of a Graphics-Based Traffic Information System Based on User Requirements", Proc. of the Human Factors Society 34th Annual Meeting, 1990.

[Gordon and Wood70] "How Drivers Locate Unfamiliar Addresses - An Experiment in Route Following", D.A. Gordon and H.C. Wood, Public Roads, Vol. 36, No. 2, June 1970.

[Gustavsson and Lindevall91] "Planning with Dynamic Parameters", R. Gustavsson and P. Lindevall, in "Final Report on Interactive Route Guidance, 1988-1991", SICS Technical Report T91:18, also presented at "Nässlingen workshop 1990", 1991.

[Gäredal et al.89] "Interactive Route Guidance - Towards a Conceptual Analysis", A. Gäredal, G. Hagert and Å. Hugosson, Manuscript, SICS-KBS laboratory, 1989.

[Gärling et al. 82] "Memory for the Spatial Layout of the Everyday Physical Environment: Different Rates of Acquisition of Different Types of Information", T. Gärling, A. Böök and N. Ergesen, Scandinavian Journal of Psychology, 23, 23-35, 1982.

[Höök and Karlgren91] "Some Principles for Route Descriptions Derived from Human Advisors", K. Höök and J. Karlgren, Proc. of the Thirteenth Annual Conf. of the Cognitive Science Society, Chicago, 1991.

[Hughes and Cole86] "What Attracts Attention When Driving?", P.K. Hughes and B.L. Cole, Ergonomics, 29, pp.377-391, 1986.

[Kay89] "Parser/Generator for Head Driven Grammars", M. Kay, First International Workshop on Parsing Technology, Pittsburgh, 1989.

[Kramer and Reichart89] "Automation and Safety Systems Engineering Aspects of Selected Prometheus Functions", U. Kramer and G. Reichart, In Second Prometheus Workshop, Stockholm, 1989.

[Kosslyn80] "Image and Mind", S.M. Kosslyn, Harvard University Press, Cambridge, Massachusetts, 1980.

[Kosslyn et al.79] "On the Demystification of Mental Imagery", S.M. Kosslyn, S. Pinker, G.E. Smith, S.P. Schwartz, The behavioral and brain science 2, pp. 535-581, 1979.

[Kuipers82] "The 'Map in the Head' Metaphor", B. Kuipers, Environment and Behaviour, 14, pp. 202-220, 1982.

[Labiale89] "Influence of In Car Navigation Map Displays on Drivers' Performance", G. Labiale, SAE Tech. Paper Series, 891683, 1989.

[Labiale90] "In-Car Road Information: Comparisons of Auditory and Visual Presentations", G. Labiale, Proc. of the Human Factors Society 34th annual meeting, 1990.

[Lindevall and Höök91] "Humans Routes - Routes for Humans", P. Lindevall and K. Höök, in "Final Prometheus Report, 1988-1991", SICS Technical Report, T91:18, 1991.

[Lynch60] "The Image of the City", K. Lynch, the M.I.T. Press, Cambridge, Massachusetts, 1960.

[Michon85] "A critical view of driver behaviour models: What do we know, what should we do?", John A. Michon, In "Human Behaviour and Traffic Safety", L. Evans and R.C. Schwing (eds), New York: Plenum Press, pp.485-520, 1985.

[Michon89] "Explanaratory pitfalls and rule-based driver models", J. A. Michon, Accident Analysis and Prevention, 21:4, pp. 341-353, 1989.

[Miller56] "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information", G.A. Miller, Psychological Review, 63, pp.81-97, 1956.

[Moore and Swartout88] "A Reactive Approach to Explanation", J.D. Moore and W.R. Swartout, Proc. of AAAI Workshop on Text Planning and Generation, St. Paul, Minnesota, 1988.

[Nwana91] "An Approach to Developing Intelligent Tutors in Mathematics", H.S. Nwana, Proc. of the 6th Int. PEG conf., Rapallo, Italy, 1991.

[Pailhous70] "La Représentation de l'Espace Urbain", J. Pailhous, Paris: Presses Universitaires de France, 1970.

[Paris88] "Tailoring Object Descriptions to a User's Level of Expertise", C. L. Paris, Computational Linguistics, 14(3), pp. 64-78, 1988.

[Parkes90] "Route Guidance Devices: Too Many, Too Soon?", A. Parkes, Proc. of the Nässlingen-90 Workshop, Available from SICS, 1990.

[Parkes and Martell90] "A Usability Analysis of a PC Based Route Planning System", A. Parkes and A. Martell, Proceedings of ergonomic society 1990 Annual Conference, E. J. Lovesay (ed.), Taylor and Francis, 1990.

[Parkes and Coleman90] "Route Guidance Systems: A Comparison of Methods of Presentating Directional Information to the Driver", A. Parkes and N. Coleman, Proceedings of ergonomic society 1990 Annual Conference, E. J. Lovesay (ed.), Taylor and Francis, 1990.

[Pylyshyn81] "The Imagery Debate: Analogue Media Versus Tacit Knowledge", Z.W. Pylyshyn, Psychological Review, 88, pp. 16-45, 1981.

[Riesbeck80] "'You Can't Miss It!': Judging the Clarity of Directions", C. K. Riesbeck, Cognitive Science 4, pp. 285-303, 1980.

[Sabey and Stoughton80] "The Drinking Road User in Great Britain", B.E. Sabey and G.C. Staugthon, TRRL Supplementary Report No. 616, 1980.

[Sadalla and Magel80] "The Perception of Traversed Distance", E.K. Sadalla and S.G. Magel, Environment and Behaviour, 12, pp. 65-79, 1980.

[Scholl and Egeth] "Cognitive Correlates of Map-Reading Ability", M. J. Scholl and H. E. Egeth, Intelligence, 6, pp.215-230, 1982.

[Simon74] "How Big Is a Chunk?", H.A. Simon, Science, vol. 183, 1974.

[Streeter et al.85] "How To Tell People Where To Go: Comparing Navigational Aids", Streeter L.A., Vitello D., and Wonsiewicz S. A., International Journal of Man-Machine Studies, 22, pp. 549-562, 1985.

[Streeter and Vitello85] "Automatic Street Route Generation: Modeling Resident Direction Givers", L. A. Streeter and D. Vitello, TM-ARH-001369, Bell Communications Research, 1985.

[Streeter and Vitello86] "A Profile of Drivers' Map-Reading Abilities", L. A. Streeter and D. Vitello, Humans Factors, 28(2), 1986.

[Thorndyke and Goldin81] "Ability differences and Cognitive Mapping Skills", P. W. Thorndyke and S. E. Goldin, RAND-N-1667-ARMY, 1981.

[Thorndyke and Hayes-Roth82] "Differences in Spatial Knowledge Acquired from Maps and Navigation", P.W. Thorndyke and Hayes-Roth, Cogn. Psychology 14, pp. 560 - 589, 1982.

[Thorndyke and Stasz80] "Individual Differences in Procedures for Knowledge Acquisition from Maps", P. W. Thorndyke and C. Stasz, Cognitive Psychology, 12, pp. 137-135, 1980.

[Treisman60] "Verbal Cues, Language and Meaning in Selective Sttention", A. M. Treisman, Quartely Journal of Experimental Psychology, 12, pp. 242-248, 1960.

[von Tomkewitsch90] "ALI-SCOUT (LISB) - a Beacon Based System for RTI and Dynamic Traffic Management", R. von Tomkewitsch, Proc. from Nässlingen Workshop, available from SICS, 1990.

[Wærn90] "Presenting Route Guidance Information: Some thought about interface design", A. Wærn, SICS technical report T91:02, 1990.

[Wærn91] "Planning Route Guidance Dialogue", A. Wærn, in "Final Report on Interactive Route Guidance, 1988-1991", SICS Technical Report T91:18, 1991.

[Wærn89] "Cognitive Aspects of Computer Supported Tasks", Y. Wærn, John Wiley & Sons, Chichester, 1989.

[Wærn and Askwall81] "On Some Sources of Meta-Comprehension", Y. Wærn and S. Askwall, Scandinavian Journal of Psychology, 22, pp.17-25, 1981.

[Wærn et al.90] "Communication Knowledge for Knowledge Communication", Y. Wærn, S. Hägglund, J. Löwgren, I. Rankin, T. Sokolnicki, A. Steinemann, LiTH-IDA-R-90-17, Linköpings University, Sweden, 1990.

[Yerkes and Dodson08] "The Relation of Strength of Stimulus to Rapidity of Habit-formation", R.M. Yerkes and J.D. Dodson, Journal of Comparative Neurology and Psychology, 18, pp. 459-482, 1908.

Chapter 9
Evaluating Adaptive Navigation Support

Kristina Höök and Martin Svensson
SICS

Since navigation in information spaces is a difficult task, it is easy to envision ways of supporting users through adapting the navigation to their knowledge, task or cognitive ability. Adaptation of navigation is, however, a "two-edged" sword since it may well just contribute to the users' workload rather than reducing it. If the space is altered, parts are hidden from the user, or links are annotated, etc. users may become confused or learn less of the structure. From the few evaluations of adaptive navigation systems that have been performed, we see an emerging pattern where depending upon the domain, only certain types of adaptive navigation works.

Furthermore, evaluations of adaptive navigation support systems fail to recognise some of the more important aspects of why certain systems provide better support than others. These studies typically measure task completion time, or how well the structure of the space is remembered. While these are among the important measurements that should be taken, other features, such as how much anxiety the system induces in users, how pleasant it is to navigate, or how much users actually learn of the information contained in the space, might be more crucial measurements.

EVALUATING ADAPTIVE NAVIGATION SUPPORT

Kristina Höök and Martin Svensson
SICS Box 1263 S-164 28 Kista Sweden
{kia, martins}@sics.se
http://www.sics.se/humle/

Since navigation in information spaces is a difficult task, it is easy to envision ways of supporting users through adapting the navigation to their knowledge, task or cognitive ability. Adaptation of navigation is, however, a "two-edged" sword since it may well just contribute to the users' workload rather than reducing it. If the space is altered, parts are hidden from the user, or links are annotated, etc. users may become confused or learn less of the structure. From the few evaluations of adaptive navigation systems that have been performed, we see an emerging pattern where depending upon the domain, only certain types of adaptive navigation works.

Furthermore, evaluations of adaptive navigation support systems fail to recognise some of the more important aspects of why certain systems provide better support than others. These studies typically measure task completion time, or how well the structure of the space is remembered. While these are among the important measurements that should be taken, other features, such as how much anxiety the system induces in users, how pleasant it is to navigate, or how much users actually learn of the information contained in the space, might be more crucial measurements.

INTRODUCTION

"Lost in hyperspace" is a feeling that is familiar to almost anyone using a computer. After a few actions, we do not know where we are, how we got there, or what our original goal was. *Adaptive navigation systems* has been proposed as a means to aid users in finding their way through information spaces. Several systems have been designed that adapts the navigation to users' knowledge (Brusilovsky and Pesin, 1994, Brusilovsky, Schwartz, and Weber, 1996, Kobsa et al., 1994, Boyle and Encarnacion, 1994), to users' preferences and goals (Kaplan et al., 1993), to users' tasks (Höök et al. 1996), or to users' spatial ability (Benyon and Murray, 1993). The hope is that if user characteristics are considered the cognitive workload can be reduced.

The question we want to pose here is exactly when those systems do in fact reduce workload? Can we find criteria that will give us better insight into how adaptive navigation should be designed to best assist users?

In order to investigate these issues further, we shall first review studies made on adaptive navigation support systems, and then outline some criteria by which these system should have been evaluated. Let us start by defining the concepts "navigation" and "information space".

Navigation and Information Space?

We use the metaphor "navigating an information space", but what do we mean by navigation and information space? In Benyon and Höök (1997) navigation is defined as both the more traditional *wayfinding* activities (when the destination is known), as well as *exploration* and *object identification*. Downs and Stea (1973) defines wayfinding as a four step process: orienting oneself in the environment, choosing the correct route, monitoring this route, and recognising that the destination has been reached. With exploration, people are not trying to get anywhere, they are not trying to find their way. Instead they are just interested to have a look around. When identifying objects, the user is not interested in the location of objects, nor is the user interested in finding a path or reaching a goal. Although object identification is somewhat akin to exploration, the purpose of the activity is different. Exploration focuses on understanding the contents of an environment and how the things are related. Object identification is concerned with finding categories and clusters of objects spread across environments, with finding interesting configurations of objects and finding out information about the objects.

In particular for the latter two kinds of activities, it is obvious that the destination is something that often is negotiable and is altered as the user moves around. Any navigation tools that fails to support renegotiations of the goal will therefore be of less use to us.

The concept of an *information space* is best understood by appealing to the notion of an information system, or information artefact. An information artefact is 'any artefact whose purpose is to allow information to be stored, retrieved, and possibly transformed' (Green and Benyon, 1996). Interactive devices such as spreadsheets, word processors, and music notations are clearly examples of information artefacts, but so are non-interactive devices like tables, documents and musical scores. What comes to mind first when discussing information spaces, are hypermedia spaces, in particular the WWW (World Wide Web). Most systems and evaluations we describe below are adaptive hypermedia systems, but we would like to make clear that adaptive navigation can also be concerned with the interface to a database system, the organisation of menus in a direct-manipulation interface, or a hierarchical file system. Despite our quite wide definition of information spaces, we shall focus mainly on adaptive hypermedia systems in the reasoning below.

ADAPTIVE NAVIGATION SYSTEMS

An adaptive system will try to infer an understanding of some user characteristics based on users' actions with the system. Based on this inferred understanding, the so called *user model*, it will then adapt its behaviour to improve the interaction with the user. Many different aspects of user characteristics can be inferred. In adaptive hypermedia[1], the *user's knowledge*, for example, is used as a basis for educational hypermedia (Brusilovsky and Pesin, 1994, Kay and Kummerfeld, 1995). The *user's familiarity with the structure of the hyperspace* is another factor that can help us limit the search for information. Vassileva (1994) uses this distinction in her adaptive navigation techniques. The *user's goal or task* is mostly used to support navigation between nodes in the hypermedia structure (Vassileva, 1994; Kaplan et al. 1993), but can also be used to decide what to show within a node (Höök et al., 1996).

As far as we know, there are no adaptive hypermedia systems that attempt to adapt to users' cognitive abilities, style, or personality traits. We believe that this is a fruitful direction of research since there are strong connections between cognitive abilities and ability to make use of hypermedia systems (Dahlbäck et al. 1996). Benyon and his colleagues (1993) made use of the fact that spatial ability together with experience of computers was related to how many errors users performed with certain database interfaces. Their system was adapted through selecting different interface for different groups of users.

---

[1] In Appendix A, we have inserted a slightly longer introduction to adaptive hypermedia.

Content Adaptation and Navigation Adaptation

Basically, there are two features of the hypermedia that can be affected by the adaptivity: the *content* of a page and the *navigation* between nodes[2]. According to Brusilovsky, we can distinguish five methods for content adaptation:

- *additional explanations* can be used for a special category of users
- *explanation variants* are used to present information in various ways depending onthe user's knowledge of the subject
- *prerequisite explanations* and *comparative explanations* change the information presented about a concept depending on the user's knowledge of other related concepts
- *sorting* means that the information pieces about a concept that are most relevant to a particular user are placed in front

Brusilovsky furthermore identifies four different adaptive techniques that affects the navigation between nodes:

- in *direct guidance* the system decides which is the next "best" node for the user to visit according to the user's goal
- in *adaptive ordering* links - on a particular page – are sorted according to the user model – the closer to the top of the list, the more relevant
- *in hiding* parts of the navigation space are hidden or restricted by removing links to non-relevant pages
- *adaptive annotation* means that we augment the links with some form of comments which can tell users more about the current state of the nodes behind the annotated links (text or visual cues)

EVALUATIONS OF ADAPTIVE NAVIGATION SYSTEMS

There are few studies of adaptive systems in general and even fewer of adaptive navigation in hypermedia systems. When first interpreting the few studies of adaptive hypermedia systems around, it can be assumed that they are, in general, quite efficient in reaching their goals. In the second of two studies of HYPERFLEX, (Kaplan et al., 1993), it was shown that the adaptive system could sometime decrease users' search time by 40%. In the study by Boyle and Encarnacion on MetaDoc, (1994), it was shown that after using the adaptive system users solved a set of reading comprehension tasks in significantly less time, and they also had significantly more correct answers. Unfortunately, other studies give a more complex view.

Edward Carter, (1996), pre-structured a hyperspace in several different ways reflecting the domain content. During the experiment, the system would switch structure when the subject turned to a new question that would be more easily solved with the other structure. Users disliked this system, and they performed worse in terms of information seeking time, time spent on each node, remembrance of the hyperspace structure after the experiment, etc. Carter speculates that this may be due to the fact that a commonly used strategy when

2 Even if this pattern is most apparent in hypermedia systems, many other applications can be analysed in similar terms. For example, navigation in a direct-manipulation interface can be viewed as navigation between states (nodes), and adaptation can either be done on how to traverse the states or on the content of each state.

users get lost in a hyperspace is to return to a "landmark" node from where they know how to proceed. If the structure is changed, they may not be able to get back to their landmark, thereby loosing their bearings.

InterBook (Brusilovsky and Schwartz, 1997), ISIS-Tutor (Brusilovsky and Pesin 1994), and ELM-ART (Weber et al., 1996) are three educational hypermedia systems where the navigation between nodes is adapted through annotating the links. In InterBook, for example, depending upon users' assumed knowledge of the domain, certain links are deemed as already known by the user (coloured in green), ready to be learnt (yellow), or too difficult (red). All three systems have been studied in order to determine their efficiency (Brusilovsky and Pesin, 1995, Brusilovsky and Pesin, forthcoming, Weber and Specht, 1997, Eklund et al., forthcoming). These studies show that the adaptive navigation support will indeed aid users in traversing the space efficiently, avoiding nodes already visited or known to the user.

In Weber and Specht's (1997) evaluation of ELM-ART II adaptive link annotation is compared to an adaptive NEXT button technique. Two measurements are taken: how many exercise pages are visited, and how many navigational steps subjects used to solve the tasks. Regarding the first measurement, it is shown that the NEXT button is of use to novices with no previous experience of hypermedia nor of the material to be learnt, while adaptive link annotation is of use to more experienced users. Regarding the second measurement (number of navigational steps), no positive effects of the adaptive annotation can be shown, in fact, subjects perform worse with the adaptive annotation. With the NEXT button, they take slightly fewer steps. Unfortunately, this study did not measure whether users learnt more with the adaptive conditions.

In studies by Brusilovsky and Pesin (1995, 1997) on the ISIS-Tutor, it was found that the adaptive system reduced the number of steps, the number of concept repetitions, and the number of task repetitions3. Two conditions were studied, adaptive annotations and adaptive hiding. The authors conclude that "a significant difference between annotation and hiding techniques of navigation support was not found, however there is an evidence that unrestricted freedom of navigation is important for the user". Comprehension time was not affected by the adaptive conditions, and the authors say that:

"Adaptive navigation support can hardly improve the quality of learning and the comprehension time, but it can reduce the number of visited nodes – thus further reducing the overall learning time."

So in both the study by Weber and Specht and the study by Brusilovsky and Pesin, the adaptive navigation support has some positive effects on the traversal of the hyperspace. But, obviously, what should be measured for an educational hypermedia system is how much more students learnt with the adaptive system, as opposed to a non-adaptive variant. Something that is not done in these two studies.

In a study by Eklund and colleagues, (forthcoming) on InterBook it was found that when using the adaptive system subjects learnt less! Eklund and colleagues explain this by:

"ANS [Adaptive Navigation Support] complicates an interface, and makes the use of the interface overall more difficult. The students had to make decisions about these annotations take them into account, and in turn this distracted them from the content. Many chose to

---

3 Tasks and concepts are presented in separate pages in this hyperspacee.

ignore it, and thus the measure of their acceptance of it is more important than whether or not it existed in the interface as they used Interbook."

So, in summary, studies of all the three educational hypermedia systems with adaptive navigation (InterBook, ELM-ART II, Isis-Tutor), show that even though they may reduce number of visited nodes, they do not promote learning – thus failing to reach their main goal. Together with Carter's findings, should we conclude that adaptive navigation fails?

We mentioned above that Boyle and Encarnacion's study of MetaDoc did show that learning was improved by the adaptive parts of the system. In the light of the studies we have just seen, this may seem strange. As it turns out, Boyle and Encarnacion do not adapt the navigation between nodes, but the information within a node. The information space is stable and does not change, and each node will contain a description of the same concept each time it is visited – what is changed is how the concept is described. Boyle and Encarnacion have taken care to make sure the different explanations obey the same pattern of description, see for example, figure 1.

General System Structure

- The AIX Operating System has three parts:
- The AIX Virtual Resource Manager (VRM)
- The AIX Operating System kernel
- The shell

General System Structure

The AIX Operating System (a group of programs that act as interface between the user and computer) has three parts:

- The AIX Virtual Resource Manager (VRM), a set of programs that manages the recources of the computer (main storage, disk storage, display stations, and printers)
- The AIX Operating System kernel, a set of program that send instructions to the VRM. It is a set of programs that control, using the VRM, the system hardware (the physical components of the system)
- A shell is often called an interface or a commmand interpreter. It is the part of the operating system that allows access to the kernel.

Figure 1. Expert and novice explanations in MetaDoc.

As we can see, stability of presentation in the two conditions (novice and expert) in MetaDoc is maintained as much as possible. The user can be sure to find approximately the same information in approximately the same format even when s/he has learnt more. The node structure is not changed. As mentioned above, Boyle and Encarnacion managed to show that their system had a significant influence on subjects' learning.

Stability of presentation seems to be a crucial factor in the success of many adaptive systems. In a study by us (Höök, 1997), we showed that our system, that adapted the content of a node, did help users to find the most relevant information in a large on-line documentation system. It also reduced the number of actions. Even more importantly, users preferred the adaptive system over the non-adaptive variant. Our system also maintains a stable interface. What is adapted within the node is which headers are "opened" and which are "closed". An opened header is one where the text under the header is shown, while a closed only shows the name of that section. By clicking on the header, the user can force a header to be closed/opened. The order of headers within a node are never restructured.

The same results are shown by Meyer (1994) and Debevc and colleagues (1996). In both systems care is taken to make the adaptivity somewhat stable and predictable. In both studies of these two systems, users' perform better and also like the adaptive system.

What is not clear in any of the studies mentioned above, is how well the system can adapt to the user? In the educational system examples, the system is supposed to keep track of the user's knowledge. But how much can be known about the user's knowledge from which links they click on? Unless we can be sure that the adaptations are "correct", these systems will be of no use. As pointed out by Judy Kay (1994) ("Lies, damned lies, and stereotypes") adaptations are always more or less stereotypical, and will therefore only be approximations of the individual users' characteristics.

Inferred design recommendations

From the few studies mentioned above we cannot claim that adaptive navigation in hypermedia is always bad, or that it is only efficient in reducing the number of nodes visited. The results may of course be due to a bad design, bad adaptations, or the fact that users are quite unused to these kinds of interfaces. Once we have standardised ways by which adaptive navigation works, users' may learn how to best utilise them. For example, on the WWW links are colored differently if the user has visited them. This very simple annotation strategy probably becomes useful to a large user population after a longer time of use. Also since this behaviour is consistent throughout the web no matter which site is visited, it can become part of users strategies in general for keeping track of where they are and where they have been.

So, further studies are needed in order to single out exactly what makes an adaptive design useful, both in the short term and in the long term. Interface design culture should not be underestimated when it comes to understanding why certain interfaces work while others do not (even when they are carefully designed!).

But what tentatively may be concluded from the studies mentioned above is that adaptations should:

- leave the interface somewhat predictable so that users do not feel lost (as in Carter's experiment),
- it should not force users to interpret advanced annotations, thus distracting them from their main tasks, and,
- finally, the adaptive navigation support should not *change* the structure of the space (as in the adaptive hiding example or in Carter's example).

Of course, all systems should be evaluated with respect to the purpose of the system, the domain, and an understanding of the intended user population. From a domain and task analysis we may see if adaptation of navigation is at all feasible. For example, in a large domain that users seldom revisit and where there is no need for the user to learn the structure of the space, adaptive guidance might be very useful. Also, in a domain where the structure is of (nearly) no importance, as for example, in a collection of movies or food recipes, where any organisation can work, adaptation as a means of structuring the space according to preferences may work really well (see e.g. Firefly).

Finally, in a domain to which users frequently return perhaps even daily, and where it is very important that they can create shortcuts through the space, adaptations based on interactions with the users might be useful. This was shown in the system created and evaluated by Kaplan and his colleagues, (1993). The system associated weights with the links between nodes. These weights were partly set initially depending upon the relation between the topics in the nodes, as well as the relation between topics and goals of users'. The users could also affect the weights themselves. The system presented a list of nodes, organised as a prioritised list that users could manipulate through dragging items higher up or lower down in the list. Users' thereby become active participants in reorganising the hyperspace.

In general, allowing users to influence the adaptations of the system are important. Preferably if users can affect the adaptation without having to understand the mechanisms behind the adaptations in detail (Höök et al. 1996, Höök, 1996, Cook and Kay, 1994).

Underlying problems

Going back to the definition of wayfinding in terms of the four steps, we find some clues to the underlying explanations to why certain adaptations will only be of limited use and may have to be combined with other methods. It seems as if the main problem with most existing adaptive navigational systems are not what they do, but rather what they do not. Considering wayfinding as described above, it is safe to say that currents systems usually address the problem of choosing the correct route. The four adaptation techniques presented by Brusilovsky: direct guidance, adaptive ordering, hiding, and adaptive annotation, all try to – in one way or the other – minimise the number of steps that a user has to take to reach a specific goal in a hypermedia system, that is, the systems focus on finding the most effective (shortest) path through the information space. However, a good adaptive hypermedia system needs to support the other steps in wayfinding as well: orienting oneself in the environment, monitoring the route, and recognising when the destination has been reached.

Firstly, the user's ability to orient herself in the hypermedia system is often reduced by the current adaptive navigation techniques, especially adaptive ordering and hiding. These techniques alter the information space either by rearranging the links or by hiding certain links from the user. The faster a user gets to know the whole of the hypermedia system, the easier it is for her to locate herself in it, thus, increasing or decreasing the space by hiding links may delay this process. Instead of choosing the shortest path through the information space, one possibility is to adapt to the most logical and easy to remember path through the system, something which may increase the user's ability to remember the structure of the space. A strategy often employed by users is to identify one node as a "landmark" node. Whenever users are lost, they "walk back" to the landmark and reorient themselves from there. If the space is adapted, this strategy will be destroyed.

Second, a space that is dynamic (changing) can reduce the ability for a user to monitor her given route. If the information space changes from time to time, for instance using hiding techniques, it will take a longer time for a user to get to know and understand the space, and thus, make it harder to monitor a route.

Lastly, adaptive navigation systems as described above, does not recognise that users may need help in reformulating their goal(s)/destination. On the contrary, the four techniques discussed above try to see to it that a user sticks to a given route. As indicated by our definition of navigation into wayfinding, exploration and identification of objects, users may not even have a particular destination in mind.

EVALUATION CRITERIA

From the studies discussed above, it also seems crucial to discuss how to measure the success of a navigational tool – what to measure in a study, and which method to employ.

Measurements

In the studies of adaptive navigation support systems above, in general, only certain measurements where taken, such as:

- number of visited nodes
- task completion time
- how well a user remembers the structure of the information space

What do these measure in terms of successful navigation in hypermedia systems? Again, it is obvious that the first and second criteria is measures of shortest paths through a hypermedia system, but is that to say that the navigation was successful? We believe that visited links and task completion time are not good measures of successful navigation, or rather, they only show one aspect of navigation. Remembering the structure of a hyperspace, can only be interesting if it is related to the structure of the domain and thus should be remembered. Alternatively, if the system is used often, and the user has to remember it in order to find relevant information.

In the introduction, we argued that navigation is a cognitively demanding activity. It increases the cognitive workload, it increases anxiety, and it put demands on users' spatial ability, and the risk is that adaptivity can create an additional burden on users. We need to find measurements on users' anxiety level, how often they feel lost, how often they feel that they have to go back in order to find their bearings, etc. We would like to put forth the following aspects:

- part of navigation is goal formulation: the goal or the quality of the goal should be influenced. In most educational systems the goals are implicit, i.e. there are several predefined goals in the form of lessons, however, in the same way that a tutor can help a student to formulate her goals, adaptive navigational systems need to be more flexible in this respect.
- what is it that we want to measure? Different domains and tasks demand for different types of measurements. For example, in educational hypermedia we would probably have to combine, for instance, task completion time with various measurements on how much of the domain that was actually learnt. Another example is a video recommending system, where task completion time could be used alone.
- anxiety is often overlooked. Even if a system adapts well users may feel out of control or disoriented, in effect, creating an element of anxiety in the user. Woods (1993) states: "flexibility, in the sense of autonomous changes by an interface mediator, creates uncertainty for the user – did something change?, why did it do that?, what will it do next?"
- navigation should be fun. A system that is fun/pleasant to use will encourage learning. It is also our belief that if users feel that navigation is fun, the risk of feeling anxious is highly reduced.

- navigation should (sometimes) engage users' curiosity. If they feel engaged and want to explore the space, they might not be as bothered about having to take many steps before reaching their destination.

Above we have discussed four aspects of navigation that is often overlooked: navigation should be fun, part of navigation is goal formulation, what do we actually want to measure, and that adaptivity introduces a certain amount of anxiety. These are all but some of the things that are interesting to measure in adaptive hypermedia systems, but they show us something important. When evaluating adaptive hypermedia systems it is not possible to only use measurements that can be automatically collected, such as number of visited nodes. In order to judge these types of systems subjective user metrics also needs to be taken into account.

Methods

Evaluating systems is a difficult task, and it becomes even more difficult when the system is adaptive. It is of crucial importance to be able to distinguish the adaptive features of the system from the general usability of the designed tool. This is probably why most studies of adaptive systems are comparisons of the system with and without adaptivity (Meyer, 1994; Boyle and Encarnacion, 1994; Brusilovsky and Pesin, 1995; Kaplan et al., 1993). The problem with those studies is obvious: the non-adaptive system may not have been designed 'optimally' for the task. At least this should be the case since adaptivity should preferably be an inherent and natural part of a system – when removed the system is not complete. Still, it is very hard to prove that it is actually the adaptivity that makes the system better unless that condition can be compared with one without adaptivity.

Another difficulty in making studies of adaptive systems is in the procedure of the study. Most adaptive systems will be really useful when they are part of the users' work for a longer period: only during that longer period can we see how the users' needs and goals vary in a "natural" way.

On the other hand, when studying adaptive navigation in real usage situation, it will be hard to single out whether it is the adaptivity that contributes to the success of the system. There are also many more variables that cannot be controlled. In the study by Eklund and colleagues the study was conducted as part of a real world teaching and learning context which means that many variables were not controlled (previous experience, strategy for interpreting the adaptive navigation, influence of the first experiment to the second, etc.).

SUMMARY

Through re-evaluating a set of studies of adaptive navigation systems, mainly within the adaptive hypermedia area, we have identified a number of weaknesses of the systems and the evaluations of them. From these we then singled out some tentative design requirements on navigational tools, in particular, for *adaptive* navigation support tools. We have furthermore suggested some other measurements by which we believe that navigational tools should be evaluated. Finally, we pointed out some difficulties attached to the methods by which adaptive navigation support systems are evaluated.

REFERENCES

Benyon, David, and Höök, Kristina (1997) Navigation in Information Spaces: supporting the individual, In S. Howard, J. Hammond, and G. Lindegaard (eds.), Human-Computer Interaction, INTERACT'97, Chapman & Hall.

Benyon, David R., and Murray, Diane (1993). *Developing Adaptive Systems to Fit Individual Aptitudes*, In W.D. Gray, W. E., Helfley and Murray, D. (eds.), Proceedings of the 1993 International Workshop on Intelligent User Interfaces, (pp. 115 – 122) Orlando, FL., New York, ACM Press.

Boyle, Craig, and Encarnacion, Antonio O. (1994). MetaDoc: An Adaptive Hypertext Reading System, *User Models and User Adapted Interaction*, (UMUAI) **4**, pp. 1 - 19.

Brusilovsky, Peter (1996). Methods and Techniques of Adaptive Hypermedia, *Journal of User Modeling and User-Adapted Interaction*, special issue on Adaptive Hypermedia, UMUAI 6, in press.

Brusilovsky, P. and Pesin, L. (1994). *ISIS-Tutor: An adaptive hypertext learning environment* In H. Ueono & V. Stefanuk (eds.), Proceedings of JCKBSE'94, Japanese-CIS Symposium on knowledge-based software engineering, Tokyo: EIC.

Brusilovsky, Peter, and Pesin, Leonid (forthcoming) Adaptive Navigation Support in Educational Hypermedia: An Evaluation of the ISIS-Tutor.

Brusilovsky, Peter, and Schwartz, Elmer (1997). User as Student: Towards an Adaptive Interface for Advanced Web-Based Applications, User Modeling, Proceedings of the Sixth International Conference, UM97, A. Jameson, C. Paris, and C. Tasso (eds.), CISM 383, Springer.

Brusilovsky, Peter, Schwartz, Elmer, and Weber, Gerhard (1996) ELM-ART: An Intelligent Tutoring System on World Wide Web, In Frasson, C., Gaulthier, G., and Lesgold, A. (eds.), Proceedings of the third international conference on intelligent tutoring systems, ITS'96, Berlin: Springer, 261-269.

Carter, Edward (1996). Quantitative Evaluation of Hypertext: Generation and Organisation Techniques, Ph.D. thesis, University of Edinburgh.

Cook, R., and Kay. J. (1994) The justified user model: a viewable, explained user model, In A. Kobsa and D. Litman (eds.), *Proceedings of the 4th International Conference on User Modeling*, pp. 73 - 78, Hyannis, Massachusetts, Mitre Corp.

Conklin, J. (1987). Hypertext: An Introduction and Survey, *IEEE Computer*, 20:17-41.

Dahlbäck, Nils, Höök, Kristina and Sjölinder, Marie (1996). *Spatial Cognition in the Mind and in the World: The case of hypermedia navigation*, The Eighteenth Annual Meeting of the Cognitive Science Society, CogSci'96, University of California, San Diego, July.

Debevc, Matjaz, Meyer, Beth, Donlagic, Dali, and Svecko, Rajko (1996) Design and Evlauation of an Adaptive Icon Toolbar, User Modelling and User-Adapted Interaction 6, 1-21, 1996.

Downs, R. and Stea, D (1973). Cognitive Representations, pp. 79-86, in R. Downs and D. Stea (eds.) Image and Environment, Chicago: Aldine.

Eklund, John, Brusilovsky, Peter, and Schwartz, Elmar (forthcoming) A study of adaptive link annotation in educational hypermedia.

Firefly, http://www.ffly.com

Green, T. R. G. and Benyon, D. R. (1996) The skull beneath the skin; Entity-relationship modelling of Information Artefacts. *International Journal of Human-Computer Studies* 44(6) 801-828.

Höök, Kristina (1996), A Glass Box Approach to Adaptive Hypermedia, Doctorate Thesis, Dept. of Computer and Systems Sciences, Stockholm University, ISBN: 91-7153-510-1, October 1996.

Höök, K. (1997), Evaluating the Utility and Usability of an Adaptive Hypermedia System, In Intelligent User Interfaces (IUI'97), Orlando, Florida, also accepted to Journal of Knowledge-Based Systems, to appear.

Höök, K., Karlgren, J., Wærn, A., Dahlbäck, N., Jansson, C-G., Karlgren, K., and Lemaire, B. (1996). A Glass Box Approach to Adaptive Hypermedia, *Journal of User Modeling and User-Adapted Interaction*, special issue on Adaptive Hypermedia, 6.

Kaplan, Craig, Justine Fenwick and James Chen. (1993). Adaptive Hypertext Navigation Based On User Goals and Context, *User Modeling and User-Adapted Interaction* **3**, pages 193-220.

Kay, Judy (1994). Lies, damned lies, and stereotypes: pragmatic approximations of users, In A. Kobsa and D. Litman (eds.), *Proceedings of the 4th International Conference on User Modeling*, pp. 73 - 78, Hyannis, Massachusetts, Mitre Corp.

Kay, Judy, and Kummerfeld, R.J. (1994). An Individual Course for the C Programming Language, *Proceedings of the Second International WWW Conference '94 Mosaic and the Web.*

Kobsa, Alfred, Müller, Dietmar, and Nill, Andreas (1994). KN-AHS: An Adaptive Hypertext Client of the User Modeling System BGP-MS, In A. Kobsa and D. Litman (eds.), *Proceedings of the 4th International Conference on User Modeling*, pp. 73 – 78, Hyannis, Massachusetts, Mitre Corp.

Meyer, B. (1994). Adaptive Performance Support: User Acceptance of a Self-Adapting System, Fourth International Conference on UM Hyannis, MA.

Vassileva, Julita (1994). User Modeling in Information Retrieval Systems, In A. Kobsa and D. Litman (eds.), *Proceedings of the 4th International Conference on User Modeling*, pp. 73 – 78, Hyannis, Massachusetts, Mitre Corp.

Weber, Gerhard, and Specht, Marcus (1997). "User Modeling and Adaptive Navigation Support in WWW-Based Tutoring Systems", User Modeling, Proceedings of the Sixth International Conference, UM97, A. Jameson, C. Paris, and C. Tasso (eds.), CISM 383, Springer.

Woods, David (1993), "The Price of Flexibility", Intelligent User Interfaces 1993, IUI'93

APPENDIX I

ADAPTIVE HYPERMEDIA

Adaptive hypermedia is a new direction of research within the area of user-adaptive systems[4]. The earliest systems date back to 1990 (Böcker at al., 1990), but most systems have been developed and described during the last three years (1993-1996). Some adaptive hypermedia systems are devoted to the tutoring situation, e.g. (Brusilovsky and Pesin, 1994). Some utilise limited variants of plan inference, e.g. our system, POP. Most use some form of simplistic explanation generation, e.g. (Kobsa et al., 1994; Kay and Kummerfeld, 1995). To some extent, these adaptive hypermedia systems take some first steps toward being personal, adaptive and multi-modal agents, although this is debatable and depends on which definition of agent and multi-modality is used.

In a special issue of Journal of User-Modelling and User-Adapted Interaction devoted to adaptive hypermedia (1996), Peter Brusilovsky has written a survey of methods and techniques for adaptive hypermedia. As he points out, hypermedia has gained ground during the last few years as a tool for user-driven access to information. In particular, the widespread use of WWW (which is hypertext based) has set a de facto standard for documentation of various kinds, and lately even allowed for more interactive systems (Rice et al., 1995).

Adaptive hypermedia marries the passive hypermedia information model with means to make systems actively adapt to the user. The systems implemented so far occupy a middle ground between user-controlled and system-controlled information retrieval.

The basic hypermedia model is quite simple. From a user perspective, all one has to do is to move between "pages" of information by following "links". Usually, the pages of information consist of text and some ready-made pictures. Following a link is done by clicking on a hotword or clicking on a hotspot in some graphics. The result of the action is (usually) a move to another page of information.

According to Brusilovsky, *adaptive* hypermedia is useful when the system is expected to be used by people with different goals and knowledge, and where the hyperspace is reasonably big. Users with different goals and knowledge may be interested in different pieces of information and may use different links for navigation. If the information space is large, an adaptive hypermedia system can help the user to search for and filter out the information most relevant to his or her needs, thereby limit the hyperspace. "Lost in hyperspace", (Conklin, 1987), has become a standard expression for what happens when the hyperspace is so large that it becomes hard to keep a model in the head of where in the structure the user is at.

As with any adaptive system, adaptive hypermedia systems may adapt their presentation of information and affect navigation based on various characteristics of the user. The user's knowledge, for example, is used as a basis for *educational hypermedia* (Brusilovsky and Pesin, 1994, Kay and Kummerfeld, 1995). In these systems the hypermedia tool supports student-driven acquisition of the learning material. The student model helps limiting the information space and aids (in particular the novice) in navigating through the material.

---

4 Though adaptive hypermedia ideas were preconceived already by Ted Nelson, (1971), who also coined the expression *stretchtext*.

The user's familiarity with the structure of the hyperspace is another factor that can help us limit the search for information. Sometimes the user can be knowledgeable in the subject area, but not familiar with the structure of the hyperspace, or vice versa; quite familiar with the structure but not with the content. Vassileva (1994) uses this distinction in her adaptive navigation techniques. Vassileva also uses the role of the user (patient, nurse or doctor) for limiting what the user is allowed to see and alter in the information space.

The user's goal or task is mostly used to support navigation between nodes in the hypermedia structure (Vassileva, 1994, 1995; Kaplan et al. 1993), but can also be used to decide what to show within a node.

As far as we know, there are no adaptive hypermedia systems that attempt to adapt to users' cognitive abilities, style or personality traits. This might potentially be a fruitful direction to explore since there are strong connections between cognitive abilities and ability to make use of hypermedia systems.

Even if WWW is what comes first to mind when we talk about hypermedia, we should remember that hypermedia can be used in several other contexts, like for example documentation of tools such as word processors. This means that the adaptive system can also observe the actions the user performs in the tool to which it is connected, and not only the actions in the hypermedia system. Given the information about the user's actions at the tool might provide a richer context to base adaptations on.

Basically, there are two features of the hypermedia which can be affected by the adaptivity: the *content* of a page and the *navigation* between nodes. These are discussed in detail below.

Content Adaptivity

Since a hypermedia system can be used by many different users with varying background knowledge, goals and needs, some of them are bound to introduce comprehension problems. Several adaptive hypermedia systems are therefore directed at attempting to provide the right level of information, and the most relevant information, (Kobsa et al., 1994; Boyle and Encarnacion, 1993; Vassileva, 1994, 1995).

Based on Brusilovsky's survey we can distinguish five methods for content adaptation:

Additional explanations

In addition to some basic presentation, a category of users can get additional information that is specially prepared for this category and that will not be shown to users of other categories.

Explanation variants

In addition to the choice of what information to present, we can add the possibility of having several different variants of the same information content. The system can then choose among these based on the user's knowledge of the subject. This is based on similar ideas as presented by Paris (1988).

Prerequisite explanations and *comparative explanations*

Both prerequisite explanations and comparative explanations change the information presented about a concept depending on the user's knowledge of other, related concepts. If some concepts must be known to the user before a certain concept is explained, those prerequisite concepts can be asserted into the explanation (Kay and Kummerfeld 1994b). A comparative explanation uses similarities between concepts to explain one by contrasting it with the another (presumably) known by the user.

Sorting

Sorting means that the information pieces about a concept that are most relevant to a particular user are placed in front.

Navigation Adaptivity

As pointed out, navigation in hypermedia can be very difficult when there are many nodes and much information in each node. Adapting to users in order to help them to navigate efficiently is of crucial importance. There are several different ways in which navigation can be effected. We need to differentiate between different kinds of links that can be available from a page. Basically, we can see three different kinds of links:

- *contextual links*: hotwords in the text / hotspots in the graphics that are placed in their context and can only be understood as part of that context.
- *index and content page links*[5]: pages that consist only of links to other pages, and are independent of the local context, but compiled from a larger hyperspace.
- *structural links*: buttons and links such as "back", "up", etc.

In order to improve navigation we can manipulate these links in various ways. Brusilovsky identifies four different adaptive techniques.

Direct guidance

In direct guidance the system decides which is the next "best" node for the user to visit according to the user's goal. This can be done on all kinds of links, but provides very limited support: the user can choose to follow the advice, or else no help will be available. Still, for novices who need a "guided tour" through the hyperspace, this might be relevant.

Adaptive ordering

In adaptive ordering we sort all the links on a particular page according to the user model – the closer to the top of the list, the more relevant. This is mostly relevant to the index and content page links or the non-contextual links. Users might have information needs that can only be satisfied through searching several information nodes in the hyperspace. Their browsing can be supported by the system (Kaplan et al., 1993; Mathé and Chen, 1994). The system can suggest which links to follow, or sort the links by their relevance to users' goals or knowledge.

Hiding

Hiding means that we hide or restrict the navigation space by removing links to non-relevant pages. The main advantage is that we do not overload users with action alternatives. The problem with this is that users might get a faulty picture of the information space as they are only allowed to view certain parts of it. Whether to apply hiding depends on the domain: hiding might be very useful in educational hypermedia.

Adaptive annotation

Adaptive annotation means that we augment the links with some form of comments which can tell users more about the current state of the nodes behind the annotated links (text or visual cues). A simple example is used in the WWW browser Netscape where visited and non-visited links have different colours. An advanced variant is to provide different annotations depending on users' knowledge or goals.

---

5 Brusilovsky (1996) differentiates between four different kinds as he divides this category ito those which are displayed as maps and those which are displayed as a list of hotwords.

Examples of Adaptive Hypermedia Systems

As we can see from this discussion, adaptive hypermedia inherits most of the problems and possibilities that adaptive systems in general provide us with. We need to find out *how* we want the system's interactions with the user to change and be adaptive, some *user characteristics* which we know influence the aspects of the system that we want to improve, and we must find computationally feasible *techniques* which realise the desired adaptive behaviour. What is different in hypermedia is the limited communication channel with the user and that hypermedia spaces have a tendency to grow to be very large. In particular, as users perceive the information space as a landscape of nodes with links connecting the nodes, the navigational problems are of crucial importance.

Let us provide some examples of adaptive hypermedia systems which illustrate some of the different content and navigation adaptations outlined above. First, the KN-AHS, (Kobsa et al., 1994), system, is an example of a system that will adapt the content of a page to the user's knowledge of the domain. KN-AHS does so through either providing more background information to a specific concept introduced in the text (for novices) or by providing more details of concept (for experts), i.e. in terms of the classification above, KN-AHS will be providing *additional explanations* and *prerequisite explanations*. KN-AHS does so through a stretchtext technique. Stretchtext enables the user or system to close or open parts of the text, like words, sentences, definitions, or paragraphs. A very similar approach was taken in MetaDoc (Boyle and Encarnacion, 1993). A useful aspect of stretchtext is that the user can always override the adaptation made by the system through clicking on a hotword to expand or collapse an explanation of the word (stretch the text).

Examples of adaptive hypermedia systems affecting navigation are HYNECOSUM, (Vassileva 1994, 1995), and HYPERFLEX, (Kaplan et al., 1993). In HYPERFLEX all pieces of information (nodes) in a hyperspace are related by weighted links. The stronger a weight is, the more relevant are the two information pieces to one another. The weights can be adjusted to the user's behaviour or to preferences of a whole group of users. This is done through machine learning: based on user feedback the system is able to adjust the weights on the links. HYPERFLEX will provide the user with a list of nodes, ordered according to decreasing relevance to the (by the user) chosen topic and goal. The user can then move a particular node in this list up or down, thereby increasing or decreasing its importance to the chosen goal. Users are also allowed to add new goals and gradually define their relations to the topics (nodes) in the database. In terms of Brusilovsky's categories (above) HYPERFLEX can be categorised as an *adaptive ordering* system.

The HYNECOSUM system and its model of users is a bit more complex than the previous examples. This is probably due to the fact that Vassileva is tackling a real-world problem. Vassileva worked together with a hospital in Münich that wanted to put all their information (such as patient journals, administration, etc.) into one big system. The information in this database cannot be static: new patients are admitted, new fever curves are entered, etc. The problem that had to be tackled was that different categories of users, like doctors, nurses, patients, are not interested in the same information and even not allowed to see the same information. In particular, they are not all allowed to update all the information. Also, within each occupation category, users will have different levels of experience of the system, and will therefore need guidance in order to find the relevant information or form to be filled. When this was done via paper-and-pencil, users would search for the form with the right physical appearance. A relational database interface previously used had therefore failed to meet the requirements as it did not display the physical appearance of the forms.

Vassileva tackled these problems through restructuring the information in a hypermedia structure where the forms were used as nodes. At the top of the information space, she placed a hierarchical task structure. She then associated the tasks with different occupation categories: so doctors will be entering diagnoses of diseases, while nurses may enter measurements of the patients' temperature. A particular person would thereby not be allowed to change a certain piece of information – they would only see certain tasks from the task hierarchy. She also inferred the user's assumed knowledge of the information space and the tasks so that an inexperienced user would only be allowed to navigate in certain restricted ways among the tasks. A more experienced user is on the other hand allowed to enter search commands that will make it possible to 'jump' to a particular piece of information or form. In this way, HYNECOSUM restricts the navigational possibilities. HYNECOSUM will also affect the information presentation by providing different presentations depending on who the user is. So, according to Brusilovsky's categories, HYNECOSUM affects navigation by *hiding* and content through *explanation variants*. It should be observed that HYNECOSUM will not infer users' tasks from their interaction with the system.

References

Boyle, Craig, and Encarnacion, Antonio O. (1994). MetaDoc: An Adaptive Hypertext Reading System, *User Models and User Adapted Interaction*, (UMUAI) **4**, pp. 1 - 19.

Brusilovsky, Peter (1996b). Adaptive Hypermedia: an Attempt to Analyze and Generalize, In P. Brusilovsky, P. Kommers, and N. Streitz (eds.) *Multimedia, Hypermedia, and Virtual Reality*, vol. 1077, pp. 288-304, Berlin: Springer-Verlag.

Brusilovsky, P. and Pesin, L. (1994). *ISIS-Tutor: An adaptive hypertext learning environment* In H. Ueono & V. Stefanuk (eds.), Proceedings of JCKBSE'94, Japanese-CIS Symposium on knowledge-based software engineering, Tokyo: EIC.

Böcker, H.-D., Hohl, H., and Schwab, T. (1990). Adept – Individualising Hypertext, In D. Diaper (ed.), *Proceedings of the IFIP TC13 Third International Conference on Human-Computer Interaction*, pp. 931 – 936, Amsterdam: North-Holland.

Conklin, J. (1987). Hypertext: An Introduction and Survey, *IEEE Computer*, 20:17-41.

Kaplan, Craig, Justine Fenwick and James Chen. (1993). Adaptive Hypertext Navigation Based On User Goals and Context, *User Modeling and User-Adapted Interaction* **3**, pages 193-220.

Kobsa, Alfred, Müller, Dietmar, and Nill, Andreas (1994). KN-AHS: An Adaptive Hypertext Client of the User Modeling System BGP-MS, In A. Kobsa and D. Litman (eds.), *Proceedings of the 4th International Conference on User Modeling*, pp. 73 – 78, Hyannis, Massachusetts, Mitre Corp.

Mathé, Nathalie, and Chen, James (1994). A User-Centred Approach to Adaptive Hypertext Based on an Information Relevance Model, In A. Kobsa and D. Litman (eds.), *Proceedings of the 4th International Conference on User Modeling*, pp. 73 – 78, Hyannis, Massachusetts, Mitre Corp.

Paris, Cécile L., (1987). Combining Discourse Strategies to Generate Descriptions to Users Along a Naive/Expert Spectrum, IJCAI'87, 1987.

Paris, Cécile L. (1988). Tailoring Object Descriptions to the User's Level of Expertise, *Computational Linguistics* **14**(3) 64-78, 1988.

Vassileva, Julita (1994). User Modeling in Information Retrieval Systems, In A. Kobsa and D. Litman (eds.), *Proceedings of the 4th International Conference on User Modeling*, pp. 73 – 78, Hyannis, Massachusetts, Mitre Corp.

Vassileva, Julita (1995). *Ensuring a Task-based Individualized Context for Information Retrieval from Hypermedia Information Systems for Hospitals*, in IJCAI'95 Workshop on Intelligent Multimedia Information Retrieval.

Chapter 10

Voices in the Forest: Sounds, Soundscapes and Interface Design

Catriona Macaulay, David Benyon and Alison Crerar
Napier University

Current work on navigation in electronic worlds is based on the assumption that geographic and electronic worlds are similar enough to make it possible to use results from work on environmental psychology and related areas in the design of electronic information spaces. The present paper is an attempt to analyze the underlying assumptions behind this approach in some detail, as well as an attempt to describe a number of different dimensions on which these spaces can differ. We also discuss how these differences might influence user behavior and design.

Voices in the Forest: Sounds, Soundscapes and Interface Design

Catriona Macaulay, David Benyon and Alison Crerar
Dept. Computer Studies, Napier University
219, Colinton Road
Edinburgh, EH14 1JD
{c.macaulay, d.benyon, a.crerar}@dcs.napier.ac.uk
http://www.dcs.napier.ac.uk/{~catriona, ~dbenyon, ~mac}

"...to you they are birds, to me they are voices in the forest"

(Feld, 1982, p.44).

The neglect of sound in mainstream interface design is striking. This can be partially accounted for by the lack of understanding about how sound is processed in the real world, and lack of inspiration about how to use sound innovatively. We present evidence from a number of sources (HCI, anthropology, film and computer games) to show that there is enough existing knowledge on sound craft to guide us, and that there exist cultures for whom hearing is the primary sense. We argue for a movement of the research agenda towards the design of soundscapes and propose a map of the soundscape together with key research questions.

1. Introduction

Consider the case of a journalist in the busy newsroom of a national daily newspaper. She spends a large part of every day looking for information on a huge range of topics. She has access to a large range of information sources, many of them electronic. When assigned a story she will be given a little information, perhaps a story from the 'wire service' that her newspaper subscribes to (wire services provide online, live-feed basic stories about breaking news events. She must now add to that as quickly as possible. A flurry of phone calls eventually dies down as she waits for people to get back to her. A good opportunity to start searching some of the available electronic sources. In the case of a fairly small, stable database searching might be relatively easy, but when searching a large, dynamic and very varied database (such as the Web, or databases of news cuttings) this may be a huge task. In such a situation the journalist may need to run several queries before feeling confident that the query has been well constructed; the process of 'query repair' (Hearst, Kopec and Brotsky, 1996).

But of course, very few journalists have the time, space or peace to devote all their attention to searching a particular information space. The newsroom is never still, people are constantly moving around. Piles of photographs, newspapers and documents are carried around the huge room in a constant stream. Well trodden paths from area to area are smoothly traversed by staff seemingly well used to the minefield of cables, papers, books, chairs and sharp edges. Impromptu conversations occasioned by a meeting on one of these paths turn into heated debates as people sitting nearby chip in. Questions, thoughts, jokes and asides fly about. As the deadline approaches, the noise and movement levels increase. Then, suddenly, a lull. The first edition has left the building. From a flurry of talking, moving around, searching text and picture databases, formal and informal meetings, visits to the library, shouting, phoning, typing, editing of photographs, and playing with layouts a newspaper has been born. Somewhere in the middle of this apparent chaos a journalist has, amongst many other things, been searching for information.

When considering the results of those early exploratory queries the journalist is looking for patterns of results. She is trying to get a sense of the data, an impression of how relevant knowledge is distributed through the system and to gather qualitative information that allows judgements to be made. In the very large, amorphous information spaces which often characterise the source of such searches, even just reading the titles of the results can be a major undertaking.

To date, graphical visualisation has been pursued as the preferred way of helping users cope with increasingly large databases. Visualisation systems such as TextTiles (Hearst, 1995) help users review the results of queries without reading text, speeding up the assimilation of patterns of results. 3D interfaces and 'virtual flyovers' of an information space might be used. However our journalist might still be faced with many screens-full of information (albeit graphical rather than textual), and a visually impaired journalist or one who just has not got the time to sit staring at a screen, would get little or no benefit at all. The time has clearly come to take the pressure off the visual channel.

The problems of living with the huge amounts of information now available to us have been the subject of much debate recently, with the phrase 'information sickness' entering the vernacular. But rather than thinking of the problem in terms of 'information overload', we can think of the problem in terms of 'word overload'. The problem is not that we have too much information, it is the way that the information is presented (typically as text). New information technologies might be prompting a shift towards a 'post-literate' culture (Escobar, 1994). We often require information systems to provide us with an impression or 'sense' of the information available, enabling us to refine our queries quickly. Whereas sight reveals surfaces, sound reveals interiors. "Sound situates man [sic] in the middle of actuality and in simultaneity, whereas vision situates man [sic] in front of things and in sequentiality" (Ong, 1967, p. 128). Might sound, then, provide a way for our knowledge worker to quickly review the results of exploratory queries from the 'inside' rather than looking at the surface from the front?

One of the most frequently cited arguments against using sound is that it would be noisy and distracting, particularly for those in the vicinity of the user. But such a view is usually based on experience or experiments with the intermittent use of isolated, foreground sounds. Brewster (1994) has commented that although sounds are often attention grabbing, they can be designed to minimise this effect. In contrast to the use of individual sounds we propose developing *soundscapes*; multiple sounds of many kinds and from many sources which exist in 3D space. We believe that soundscapes, if well designed and integrated with applications, can provide the sort of interaction which visualisation cannot (at least by itself).

The paper is organised as follows. In section 2, we review existing uses of sound at the interface. We believe that the main reason that sound has been under-exploited in computer systems is because of the western bias towards the visual channel and a narrow psychoacoustic approach to the psychology of audition. Section 3 examines the psychological and anthropological basis for understanding and interpreting sounds. Section 4 presents a framework for understanding soundscapes and section 5 discusses how designers might create soundscapes. A brief conclusion is provided in section 6.

2. Approaches to Using Sound in HCI

Despite its potential as an input or output modality, speech has not gained wide acceptance in HCI. The problems surrounding natural language processing of speech input are well-known and speech output is still limited to the artificial voices of text to speech converters. Non-speech sound has been used in human-computer interface design in two main ways. Auralisation is concerned with presentation of information in sound and has been used to facilitate monitoring of systems events and processes. Sonification has been used to provide the auditory equivalent of the visualisation of data set views, mapping numerically represented relationships through sound, usually abstract or musical tones (Kramer, 1994). Sonification has been used in systems for analysing financial, scientific and geographic data. Examples of auralisation systems include *ShareMon* (Cohen, 1994), a network activity monitoring system, and the *Audible Web* (Albers and Bergman, 1995) which provided feedback and monitoring information for the *Mosaic* web browser.

Research in the area of using non-speech sound in auralisation systems has concentrated on two main sound types: earcons (sounds based on abstract tones) and auditory icons (which are based on everyday, or environmental sounds. *SharedARK* (Gaver, 1994) is a system which allows users to monitor and ensure the smooth operation of a soft-drink plant assembly line model. It was observed that *SharedARK* helped users hear the plant as "an integrated, complex process" (Gaver, 1994). Earcons, on the other hand, are "abstract, synthetic tones that can be used in structured combinations to create sound messages to represent parts of an interface" (Brewster, Wright and Edwards, 1994, p. 473). In the context of a reproduced sound on a computer system, however, the distinction between earcons and auditory icons becomes less clear cut. What might clearly be the sound of two hands clapping when we are looking at the person doing it, may simply be some abstract tones in a particular rhythm without this visual confirmation (Brewster, 1994).

The use of music in auditory displays has been used to visualise computer algorithms (Alty, 1995). Music was used in *CAITLIN* to 'visualise' Turbo Pascal programs as an aid to novice programmers in the debugging process (Vickers and Alty, 1996). Alty's work is primarily concerned with a classical approach to music. However many sounds in the environment could be regarded as musical: bird song, rhythmic crashing of waves against the shore, etc. The sounds emanating from some sonification systems can be musical, and researchers in the field have commented on the possibility of composing representations of multivariate data sets using structures informed by musical principles (Mayer-Kress, Bargar and Choi, 1994). However this is not the place for a discussion of exactly what music is. In terms of interface design, perhaps it would be better to regard music as a continuum, from the incidentally melodic, harmonic or rhythmic noises that emanate from the environment, to elaborate human creations conforming to various predetermined musical conventions.

Sound has been most successfully exploited in interface design through multimedia games such as ZORK NEMESIS (Activision) and Myst (Broderbund Software). ZORK makes extensive use of environmental sounds to help the user explore this interactive mystery. Sounds are used to give feedback about actions taken, and ambient noises supply information about the kind of environment. Sound is also used to help players navigate about the complicated landscape. One section finds the player in a ruined temple with a large quadrangle containing two fountains and many rooms off each of the walkways. The sound of the fountains gets louder as the quadrangle is approached, and is heard in a roughly appropriate 3D location 'around' the players heads depending on where they 'are' in relation to the fountains. Each fountain has a different water sound, if the player stands between them and 'turns' around 180 degrees the sounds also 'move'. Sound is also used to inform players of events happening 'off-screen'.

In addition to environmental sounds, Myst uses music to signal the emotional tone of a particular area of the game's landscape, as well as using atmospheric noises in a musical way (such as the sound of the waves on the shore at the start of the game). Music is a good example of the use of sound to communicate complicated ideas and concepts, not just simple facts and figures. Apart from multimedia games, music is used extensively in films. In Western society in particular, a great many real-life soundscapes feature music from many sources - the car radio, the TV in the other room, the little boy upstairs torturing a violin. From a pragmatic point of view, our familiarity with music and our ability to remember vast amounts of it over long periods of time (Deutsch and Pierce, 1992) would indicate that this is a sound type worthy of more detailed investigation by HCI designers.

One system which seeks to integrate sounds into a soundscape is *loudSPIRE* (LoPresti and Harris, 1996), an auditory display developed to supplement a visualisation system for analysing information in large document-based databases. Inspired by the idea of developing a display like that of everyday acoustic environments, a display schema was developed using different kinds of sound (everyday, abstract and music). Preliminary feedback suggests that users found the system not only usable, but pleasant. This would point towards what might be one of the most important benefits of using a soundscape approach to auditory display design — such systems are likely to be aesthetically pleasing and enjoyable to use.

3. Understanding sound

Much research on audition has been based around laboratory studies of abstract or musical sounds and with how we attend to the primitive stimuli of pitch, loudness, timbre, etc. (see for example Pitt and Crowder, 1992; Marks, 1993). The traditional model of sound perception is based on an information processing framework, which sees audition as the internal analysis and transformation of raw sensory information (McAdams, 1993). Work on how people process everyday sounds is less well developed. Gaver proposes that in everyday listening we are primarily attending to the events that caused the sound (Gaver, 1993). Inspired by Gibson's ecological approach to vision, Gaver argues that the perception of everyday sounds is usually of complex events and entities, unmediated by inference and memory.

More recently, Truax has developed an inter-disciplinary model of auditory perception, based on communication studies, which he calls acoustic communication (Truax, 1996). Concerned primarily with environmental sound, the model is based on the idea of *information exchange*, with the listener at the centre of the process. He proposes that sound mediates the relation of the listener to the environment and that treating sounds in the abstract misses the rich variety of personal and cultural meanings of sound. The acoustic communication approach highlights the importance of the listener's context and prior experience, providing a richer view of the process of audition than can be found in the traditional information processing model.

Broadening and enriching our views of sound perception is central to the effective use of sound in human-computer interfaces. Non-speech sound tends still to be regarded as a secondary communication channel which can convey only the simplest of messages. The visual bias of western cultures has been demonstrated by a number of anthropological studies. For example, studying the importance of birds to the Kaluli of Papua New Guinea, it was observed that "When asked direct questions that include the name of a bird, the response "It sounds like X" is universally presented by a Kaluli before any sort of "It looks like X" statement." (Feld, 1982, p. 72). To the Kaluli the primary medium through which the world is conceptualised and related to is audition, not vision.

Sound can also be used to convey rich information and to enhance narrative. A study of the Kalapalo of north-eastern Brazil concluded that "Through sound symbols, ideas about relationships, activities, causalities, processes, goals, consequences and states of mind are conceived, represented and rendered apparent to the world." (Basso, 1985, p.311). Writing about the Kalapalo's narrative performance art, Basso comments on the striking weaving of large numbers of sounds with the more straightforward narrative text of the story; the sounds help the audience follow the story. In stories birds can appear either as birds or as representative of magical beings. The sound type used when the bird appears lets the audience know in which capacity the entity should be read. Anthropological studies also provide us with attempts to understand the impact of contextual/cultural factors on listening (Cole, 1996).

From such studies it can be seen that both how we hear and, most importantly, how we derive meaning from what we hear is situated in a cultural and historical context. This concern is mirrored in the ecological psychology and acoustic communication approaches, which highlight the importance of considering the real acoustic environment we live in; that complicated balance of multi-layered sounds that makes up the soundscape. Interestingly, this concern with the context of communication and cognition is currently receiving a lot of attention in HCI research through attempts to incorporate new ideas such as situated action, distributed cognition and activity theory (e.g. Nardi, 1996).

4. From Sound to Soundscapes

Given the rich and varied views of sound in different societies, trying to understand and analyse soundscapes is challenging. How do we chunk the myriad sound waves that compose an everyday soundscape? Where does one sound end and another begin? As Schwartz (1974) has pointed out, we cannot 'freeze frame' sound as the result would be silence. We do not experience sounds as a continuum but as bits of sound which we arrange into meaningful units. Nevertheless, a more detailed description, a 'map' if you like, of soundscapes is necessary if we are to design and develop soundscapes. This map must provide a means whereby the designer can analyse real-life soundscapes (to improve understanding of the way people interact with soundscapes and generate meaning from them) and a framework for the creation of artificial soundscapes.

The poverty of individual sounds has been demonstrated by Macaulay (Macaulay, 1996). This study showed only a moderate positive correlation between how easy it was for users to identify the source of particular everyday sounds and their preferences in mapping those sounds onto system features. The study also showed that laboratory-based studies of everyday, situated sounds were an inappropriate way to understand their potential use in HCI. One needs to use models and techniques from a range of disciplines to understand and describe the soundscape. Using a Gibsonian model, Gaver has produced a classification of sound-producing events identifying classes of materials and the interactions that can cause them to sound (Gaver, 1993). In his theory of acoustic communication Truax (1996) gives centre stage to information, although this time situating the listener's subjective interpretations of the soundscape more clearly in the middle of the acoustic experience.

Ferrington (1996) has suggested that real life soundscapes are composed of three layers of acoustical information: foreground sounds, contextual sounds which support the foreground sound, and background or ambient sounds. Listen to the sounds around you now. Some, such as the whine of a disc drive will be background sounds, providing reassurance or information about the state of the world. Others provide context, helping you orient to the nature of the environment. Foreground sounds such as the beep of the computer attract your attention. Isolated, intermittent sounds (as often found in interface designs) only provide a single layer of acoustical information, that of the foreground. Sound synthesising algorithms such as those developed by Gaver (1994), and sound spatialization technologies attempt to generate direction, distance or environmental context cues for individual sounds. However using Ferrington's model we can see that without additional contextual and background sounds, deriving meaning from isolated sounds is compromised.

We now turn our attention to the kinds of information that we might be able to communicate in an artificially created soundscape built using this acoustical layers/sound types framework.

	Category	Example
1	Visible entities and events	The phone ringing
2	Hidden entities and events	The photocopier round the corner being used
3	Imagined entities and events	Something big is happening on the political desk (it has gone quiet).
4	Patterns of events/entities	Someone is batch copying a large document
5	The passing of time	It's nearly deadline time (because the shift change is happening)
6	Emotions	The sports desk sub-editor is unhappy (tapping)
7	Position in Euclidean/acoustic space of entities/events and of the listener	The editor is at the foreign desk behind me (can hear his voice)

Table 1 - Information Categories

Let us return to the newsroom. The journalist has almost finished her story, although she is waiting for a call from a source. The hall is full of people as a shift change takes place. Everyone is busy, except for the sports sub-editor who has been held up as he waits for a print, and is tapping a ruler on his desk. The editor is wandering around asking people what they are doing.The soundscape might supply the journalist with information about:

We can now propose a map of the soundscape built on the three dimensions of - sound type, acoustical information level and information category, and with the listener situated in an environmental/cultural/historical context (see figure 1). To use a film metaphor for a moment, it is possible to conceptualise the soundscape as a scene in a film (situation/context) during which the various 'characters' (sounds) interact in various 'places' (acoustical information levels) via particular 'actors' (sound types) to impart various elements of a story (information categories). The audience members use their various past experiences, familiarities with the story/conventions, etc. to inform their reading of the scene. A filmic metaphor has the additional advantage of capturing the temporal nature of the soundscape, since as we have already noted, sounds only exist in their unfolding over time, just as films only exist when the projector is turning. Figure 1 provides a two-dimensional representation of this conceptualisation. The sound types may appear at any of the three acoustic information levels and give rise to any of the seven information categories. In Figure 2 a 3D rendering of this is provided to try to give the impression of the user situated, surrounded by the various combinations of the soundscape dimensions. We recognise the limitations of these visual representations, but trying to capture the nature of sound in a visual way is indeed challenging.

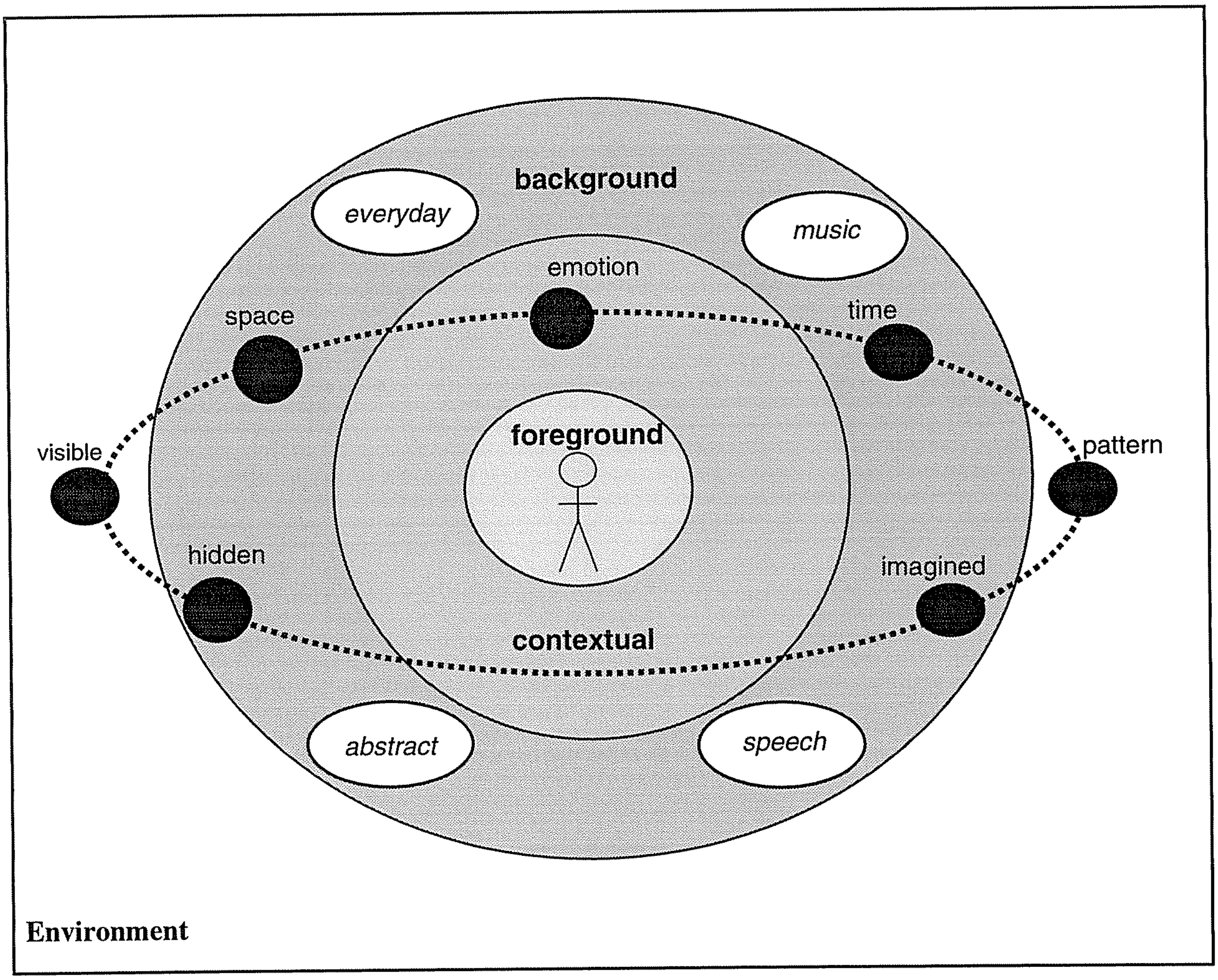


Figure 1: Map of the Soundscape

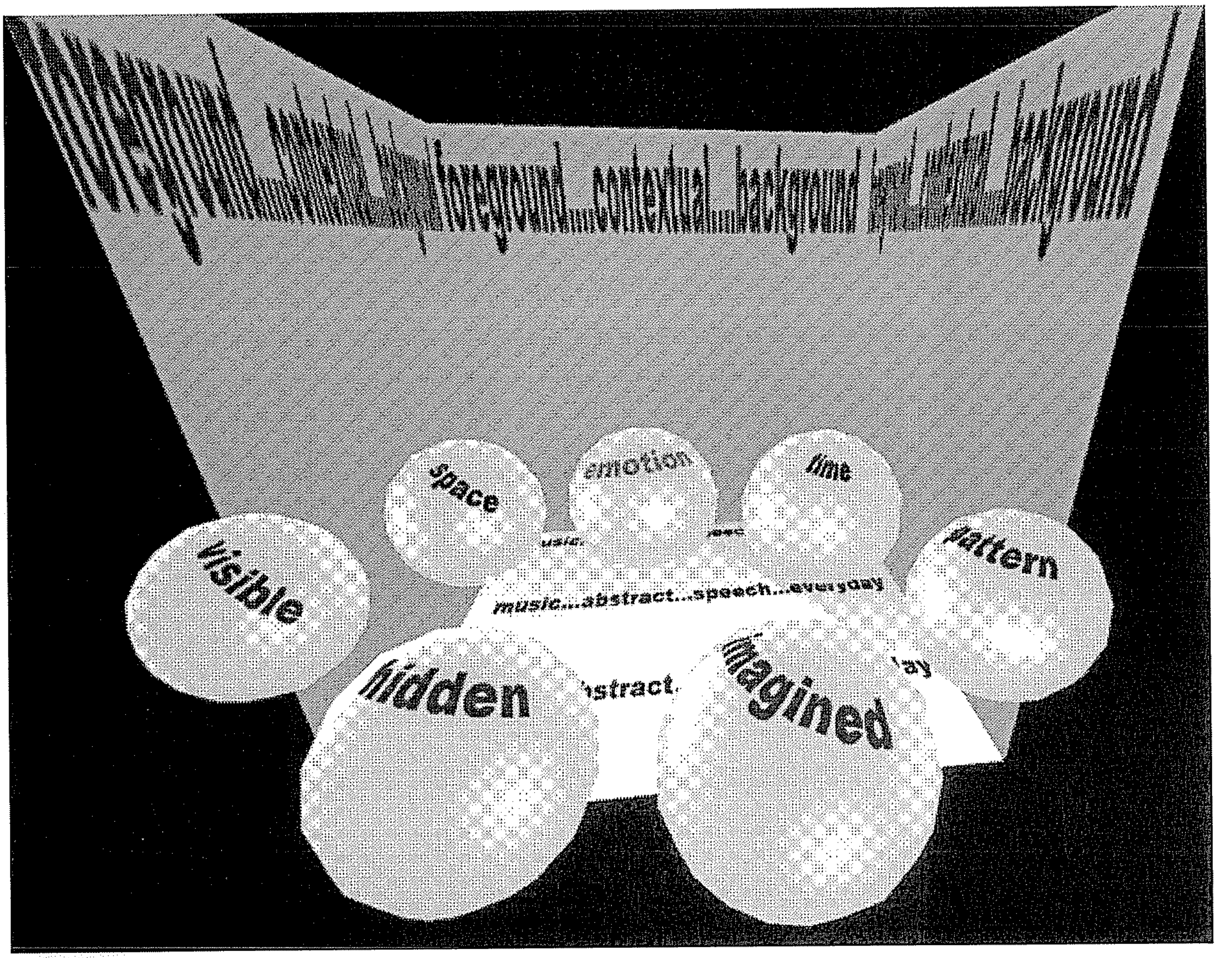


Figure 2: 3D rendering of map of the Soundscape

Whilst the parts of the soundscape are described at a fairly high level, they are sufficiently detailed to allow us to develop an idea of the main elements of the soundscape. If you were to stop for a moment and listen to the soundscape you are in, we suggest that you would be listening to different types of sound at different levels, and that these sounds and the interactions between them would provide you with information that would map onto the categories listed above. When approaching the design of individual elements in an artificial soundscape a framework such as Gaver's (see above) may be turned to for an appropriately lower level of detail. However, for our purposes this model provides a useful framework for investigating the way people hear and interpret the soundscapes, not just the sounds, around them.

5. Designing Soundscapes

If designers are to take advantage of the soundscape concept, then they must be willing to leave aside their visual bias. Designers must consider how the various information entities described above can best be presented and integrated. We believe that by taking a holistic view of the design and by considering the narrative presented by the soundscape, impressive interfaces to applications can be developed. Brenda Laurel has persuasively argued that interface design and theatre share many similarities, stemming from the fact that both attempt to represent real life, and that whilst these representations are not the same as real life they have real life results (Laurel, 1993). Rather than look to theatre for inspiration for soundscape design, we prefer to look to an area which has clearly exploited sound effectively; film.

Outlining the key techniques involved in sound editing, Reisz and Millar (1968) begin by noting the importance of taking into account the way in which we hear sound in the everyday world. Although they remark that simply copying the way sounds appear in real life would result in a cacophony, they emphasise one cannot simply 'cut' from one sound to another since individual sounds are part of a soundscape and rarely suddenly appear and disappear - more often sounds fade in and out of our attention in an ever changing acoustic environment. Emphasising the importance of sound in establishing a sense of place and orienting the user within it, they note that asynchronous sound is even more important than synchronous sound since it can add information about things we cannot see on screen, allowing the filmmaker to tell a more complex 'story' in a scene than might otherwise be possible with visuals alone. Monaco (1981) has commented that sound effects do the real work of creating the sound environment that is so important in bringing a scene to life and establishing a sense of place. Reisz and Millar (1968) also point out how sound allowed for new ways of determining pace and suggesting the passing of time.

Sound is often used to direct the audience's attention to events not currently in the foreground but which the director wants them to attend to. Visconti made dramatic use of this kind of 'attention directing' technique in the film Ossessione (Italy, 1942), which tells the story of lovers who kill the woman's husband. We see the lovers meeting in a public square, but as they approach each other the camera pulls away to visually foreground a group of children playing. As we watch the children play in the visual foreground, the acoustic foreground is occupied by the sound of the lovers arguing about what to do (they are by now only barely visible in the background). Contrasting the foreground visual with the 'foreground sound' serves to concentrate our attention on the dialogue more tightly than had the characters themselves been in the foreground shot. Silent filmmakers had to rely on all manner of visual conventions and stylistic tricks to convey what could be communicated far more simply with a line of dialogue, a snatch of music or a sound effect. Supporting the proposal that sound altered the nature of filmmaking, Neale (1985) comments that the binding of sound and image "unifies the filmic text, the space of the auditorium and the psyche of the viewer".

here are clearly many similarities between the situations in which sound is used in the cinema and those in which it could be used in HCI. Table 1 presents these in relation to the soundscape map presented in Figure 1. Films typically make use of all four sound types, and sound is presented in all

Information Category	Film	HCI
1, 2, 3, 4, 5	Provide auditory confirmation that something has happened.	Provide feedback about actions taken.
6	Signal the 'emotional tone' of a scene.	Generate affective responses in the user (e.g. re-assurance, alarms).
2, 3, 4	Let us know of something happening 'off screen'.	Inform user of something happening elsewhere on a network or in an occluded area.
1, 4, 5, 7	Draw attention to an event, direct attention to a new part of the scene.	Inform user about an event in a part of the screen/another window which is not currently the focus of attention.
1, 7	Distract attention whilst a change is made in the visual scene which might disrupt our sense of a continuous narrative.	Distract the user from screen changes.
1, 2, 3, 4, 5, 6, 7	Alert the audience to something that is about to happen (e.g. ominous music, etc.)	Get the user's attention.
7	Help the audience orient to a new location.	Help the user navigate in 'virtual' environments.

Table 2 - Similarities between Use of Sound in Film and HCI (see Section 4 for information categories)

three acoustical information levels. Most film soundtracks would therefore fit the map of a soundscape presented in this paper.

Film makers have learned a lot about how to exploit sound in their art. A soundtrack is not just 'real-world sound'. It is carefully constructed to sound like a real-world soundscape so that it is, in general, unintrusive. Equally it is designed to inform and evoke responses in the audience and to focus attention on significant occurrences. Laurel (1993) has observed that HCI is really about designing imaginary worlds. The story which the designer is trying to tell can be significantly enhanced by the appropriate use of sound.

6. Conclusions and Future Work

Although the auditory channel of communication has a potentially important role to play in human-computer interfaces, the intermittent use of isolated sounds in interface design has often proved to be ineffective in terms of usefulness and user acceptance. If we look at the way we perceive and interact with sounds in the real world, where individual sounds exist in rich and complex soundscapes, we see a different story. Given the incompleteness of the currently available models of audition, interface designers seeking to use everyday sounds in their work can learn much from filmmakers and other cultures. One of the main lessons from cinema is not so much that filmmakers make 'good sounds' but that filmmakers use sound in good ways and integrate it well with the visual elements of their films. The challenge for interface designers is as much that of finding good ways to use sound in computing as it is in designing good individual sound elements (Kramer, 1994).

We have presented a variety of scenarios illustrating the challenges faced by designers of interfaces to information retrieval and information exploration interfaces. These scenarios, derived from our own ongoing year-long ethnographic study of information seeking at a national daily newspaper in the UK, highlight the complex interactional situations many users find themselves in. Few interface designers today can pretend that 'users' use their systems for clear-cut reasons, over sustained periods of time, on a regular basis and with the peace and quiet required for highly attentive interaction. We have suggested that the use of sound, and soundscapes, offers many solutions to the overwhelming of the GUI. To this end, we have presented a map of the soundscape which can be applied equally to real-life and artificially generated soundscapes. The research agenda for developing soundscapes needs to focus on:

- should particular sounds be mapped to particular system events or data?
- how important is it to provide a 3D soundscape and how should this be achieved?
- how immersive should the soundscape experience be?
- what role will affect have on users' ability to work with soundscapes?
- how much training will users from visually biased cultures require/accept?
- how do we compose a soundscape?
- how should sound elements be created?

The arrival of sound technology fundamentally changed the way filmmakers thought about and constructed films (Neale, 1985). In time, the arrival of sound in the interface may have the same far-reaching effects. At the very least by considering the place of sound in interface design an opportunity arises to consider again existing design paradigms and practices. Work on using sound in the interface can inform the broader debates taking place in the HCI community about how we model users and tasks, define requirements, develop design methodologies and evaluate our work.

Rainforests are a powerful metaphor for large, complex information spaces — they are dense, contain huge number of entities which cannot be seen, are easy to get lost in and can only satisfactorily be visually perceived from a great distance. This paper opened with a quotation from Feld who, having realised the importance of birds in Kaluli society, decided to try and classify the many species in the area. Working with the Kaluli he used a Western approach to classification, trying to record the names and physical descriptions of each species. One of the villagers, tiring of the exercise, explained that the problem that to Feld they were birds, but to him they were 'voices in the forest'.

The paradigm shift demanded by the Kaluli villager is analogous to the change of thought needed in interface design. Perhaps it is because we live in such a visually-biased culture that software designers appear afraid of using sound at the interface and seem shy about looking to games and film for inspiration. Now that the desk-top computer is but a small window onto a vast, diverse, dynamic information space, the visual channel on its own is too narrow. We can no longer easily see all the available information, but perhaps we can hear it.

REFERENCES

Albers, M.C. 1996. Auditory Cues for Browsing, Surfing, and Navigating. In *Proceedings of ICAD 96*. Available via Internet (visited 20/1/98): http://www.santafe.edu/~icad/ICAD96/proc96/albers5.htm

Alty, J. L. 1995. Can We Use Music in Human-Computer Communication? In. Kirby, M.A.R., Dix, A.J. and Finlay, J. E. (eds.). *People and Computers X. Proceedings of HCI '95*. Huddersfield: UK. Cambridge University Press. 409-423.

Basso, E. B. 1985. *A Musical View of the Universe*. Philadelphia: University of Pennsylvania Press.

Brewster, S. A. 1994. *Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interface*. PhD Thesis. University of York.

Brewster, S.A., Wright, P.C. and Edwards, A.D.N. 1994. A Detailed Investigation into the Effectiveness of Earcons. In *Auditory Display; Signification, Audification, and Auditory Interfaces*. Kramer, G. (ed.). Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Addison-Wesley. 471-498.

Cohen, J. 1994. Monitoring Background Activities. In *Auditory Display; Signification, Audification, and Auditory Interfaces*. Kramer, G. (ed.). Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading MA: Addison-Wesley. 499-532.

Cole, M. 1996. *Cultural Psychology*. :Harvard University Press.

Deutsch, D. and Pierce, J. R. 1992. The Climate of Auditory Image and Music. In *Auditory Imagery*. Reisberg, D. (ed.). Hillsdale: Lawrence Erlbaum Associates.

Escobar, A. 1994. Welcome to Cyberia: Notes on the Anthropology of Cyberculture. *Current Anthropology*, **35(3)**, 211-231.

Feld, S. 1982. *Sound and Sentiment: Birds, Weeping, Poetics and Song in Kaluli Expression*. Philadelphia: University of Pennsylvania Press.

Ferrington, G. 1994. Keep Your Ear-Lids Open. *Journal of Visual Literacy*

Gaver, W.W. 1993. What in the World do we Hear?: An Ecological Approach to Auditory Event Perception.. *Ecological Psychology*, **5(1)**, 1-29.

Gaver, W.W. 1994. Using and Creating Auditory Icons. In *Auditory Display; Signification, Audification, and Auditory Interfaces*. Kramer, G. (ed.). Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading MA: Addison-Wesley. 417-446.

Hearst, M. 1995. Visualisation of Term Distribution Information in Full Text Information Access. In *Human Factors in Computing Systems* Proceedings, Annual Conference Series, 1995, ACM SIGCHI. 59-66.

Hearst, M., Kopec, G. and Brotsky, D. 1996. Research in Support of Digital Libraries at Xerox PARC Part II: Paper and Digital Documents. *D-Lib Magazine*. June. Available via Internet (visited 20/1/98): http://www.dlib.org/dlib/june96/06contents.html

Kramer, G. 1994. An Introduction to Auditory Display. In *Auditory Display; Signification, Audification, and Auditory Interfaces*. Kramer, G. (ed.). Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading MA: Addison-Wesley. 1-78.

Laurel, B. 1993. *Computers as Theatre*. Reading, MA: Addison-Wesley.

LoPresti, E. and Harris, W. M. 1996. loudSPIRE, an Auditory Display Schema for the SPIRE System. In *Proceedings of ICAD 96.* Available via Internet (visited 20/1/98): http://www.santafe.edu/~icad/ICAD96/proc96/lopresti.htm

McAdams, S. 1993. Recognition of Sound Sources and Events. In *Thinking in Sound: The Cognitive Psychology of Human Audition.* McAdams, S and Bigand, E. (eds.). Oxford: Clarendon Press. 146-198.

Macaulay, C. 1996. *Cinematography and Interface Design: Lessons in the Use of Sound (An Investigation into the Use of Everyday Sounds in Interface Design).* Unpublished M.Sc. Dissertation. Napier University

Marks, L. E. 1993. Contextual Processing of Multidimensional and Unidimensional Auditory Stimuli. *Journal of Experimental Psychology: Human Perception and Performance.* **19(2)**, 227-249.

Mayer-Kress, G., Bargar, R. and Choi, I. 1994. Musical Structures in Data from Chaotic Attractors. In *Auditory Display; Signification, Audification, and Auditory Interfaces.* Kramer, G. (ed.). Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading MA: Addison-Wesley. 341-368.

Monaco, J. 1981. *How To Read A Film: The Art, Technology, Language, History and Theory of Film and Media.* New York: Oxford University Press.

Nardi, B.A. 1996. Studying Context: A Comparison of Activity Theory, Situated Action Models, and Distributed Cognition. In *Context and Consciousness: Activity Theory and Human-Computer Interaction.* Nardi, B.A. (ed.). Cambridge, MA: The MIT Press, 69-102.

Neale, S. 1985. *Cinema and Technology: Image, Sound, Colour.* London: Macmillan Education. .

Newell, A. F. 1995. Extra-Ordinary Human-Computer Interaction. In *Extra-Ordinary Human-Computer Interaction: Interfaces for Users with Disabilities.* Edwards, A.D.N. (ed.). Cambridge: UK. Cambridge University Press.

Ong, W. J. 1967. *The Presence of the Word: Some Prologomena for Cultural and Religious History.* New Haven: Yale University Press.

Ong, W. J. 1982. *Orality and Literacy. The Technologizing of the Word.* London: Routledge.

Pitt and Crowder. 1992. The Role of Spectral and Dynamic Cues in Imagery for Musical Timbre. *Journal of Experimental Psychology: Human Perception and Performance.* **18(3)**, 728-738.

Reisz, K. and Millar, G. 1968. *The Technique of Film Editing.* London. Focal Press.

Schwartz, T. 1974. *The Responsive Chord.* New York: Anchor Books.

Truax, B. 1984. *Acoustic Communication.* Norwood, NJ: Ablex Publishing Corporation.

Vickers, P. and Alty, J. L. 1996. CAITLIN: A Musical Program Auralisation Tool to Assist Novice Programmers with Debugging. In *Proceedings of ICAD 96.* Available via Internet (visited 20/1/98): http://www.santafe.edu/~icad/ICAD96/proc96/papers1.htm

Chapter 11

A Comparative Study of Digital and Cinematic Space with Special Focus on Navigational Issues

Per Persson

SICS

In order to better understand the characteristics of digital space, this paper investigates another artificial environment - cinema. By presenting some very common visualization techniques employed in mainstream film making, the paper first tries to compare the visualization techniques in the two media and see if the cinematic ones has any equivalents in computer interfaces. Secondly, it speculates around the possibility of implementing the cinematic techniques in interface design. In connection to this many parameters have to be considered. Digital and cinematic spaces are used for different purposes. Digital environments are much more diverse than cinematic ones (abstract-realistic). Throughout, the discussion aspires to relate to the issue of navigation and in what way cinematic visualization techniques might support the user in this respect.

A Comparative Study of Digital and Cinematic Space with Special Focus on Navigational Issues

Per Persson[1]
SICS Box 1263 S-164 Kista Sweden
per.persson@mail.film.su.se
http://film-server.film.su.se/Perhome/Index.html

ABSTRACT
In order to better understand the characteristics of digital space, this paper investigates another artificial environment - cinema. By presenting some very common visualization techniques employed in mainstream film making, the paper first tries to compare the visualization techniques in the two media and see if the cinematic ones has any equivalents in computer interfaces. Secondly, it speculates around the possibility of implementing the cinematic techniques in interface design. In connection to this many parameters have to be considered. Digital and cinematic spaces are used for different purposes. Digital environments are much more diverse than cinematic ones (abstract-realistic). Throughout, the discussion aspires to relate to the issue of navigation and in what way cinematic visualization techniques might support the user in this respect.

1. INTRODUCTION

As computer interfaces become more like graphics and images, the interest in visual arts has increased among designers. As interfaces has turned from punch cards, through keyboards, to graphical screens with mice to manipulate the increasingly advanced images, the visualization techniques used in graphics, painting, photography and cinema is coming into focus. This paper seeks to compare computer space with *cinematic* space, with special emphasis on navigational issues. Why is space in most mainstream films experienced as clear and coherent, while computer users customarily get lost in hyperspace? Can computer people learn something from the study of film? Or are the two spaces different and incompatible?

Let us start with a general level and two apparent differences between the two and see what they will lead to.

First and most obvious, cinematic space is not interactive in the way computer space is. Games, data bases, MUDs, hypermedia, the web, word processing programs and other kind of software, presents objects and space that the user is able to manipulate, rearrange and present in an optional order. In word- or graphic-software space, the user creates and changes objects like documents and images. In hypermedia the user more or less chooses the order in which the information is presented.

In cinema nothing of this seems to be the case. The spectator can not affect cinematic objects or warn or punish characters in cinematic space (even though she would like to). The viewer cannot change the course of events. And the order in which story information is presented is determined by the filmmakers, by the *linearity* of the cinematic medium. Although interpreting a film always includes a *mental* jumping back and forth from 'now' and backwards to different points in the film, and although the spectator have to *mentally* rearrange the order in, for instance, a flashback, the cinematic 'text' *itself* cannot be reordered or changed. This is, however, not to be seen as a deficiency visavi computer space. In order to achieve narrative and aesthetic effects (suspense, surprise, humor, tragedy etc.), it is often important for the makers of the film to present the right kind of information at precisely the right moment in the linear chain. This is the reason why it is difficult to generate effective hypertext narratives (Plowman, 1996).

Secondly, in contrast to films, computers are mostly used as *tools*, and computer space is designed with the purpose of maximizing this feature (by inhabiting it with intuitive objects and buttons). Computer interfaces are spatialized mainly in order for the user to solve some problem or execute a task more efficiently, whether this be word-processing, layout, information seeking etc.. Cinematic space on the other hand is inhabited with objects, characters and actions, not to solve some problem, but mostly for the *pleasure* of it. We do not make the effort to go to the cinema (or go to the video store) because we have a task to be solved, but because we want to be entertained, surprised, frightened, excited or whatever. In mainstream fiction cinema, for instance, the cinematic space is used for *narrative* purposes.

However, even though computer and cinematic space are used for different purposes in mind

[1] Also affiliated with Department of Cinema Studies, University of Stockholm, Sweden.

(tool-pleasure), this is not inherently necessary. Computer games or MUDs are cardinal example of computer space used for pleasure. And documentaries and instruction films are being watched because the spectator wants to learn something, i.e. used as a tool. In this way I think the apparent discrepancy between cinema and computer space is not about the space itself, but the way the space is *used* by the industries. The purpose of moving images was in the beginning of this century heatedly debated (science, documentaries, entertainment, poetry or what?), but it was the entertainment industry that picked it up most vividly and made it into the pleasure product we normally associate with mainstream cinema (Gunning, 1994). Computers on the other hand was from the beginning in the hands of scientists and their view of what they should be used for has continued to be the dominating conception. With computer games, the web and hypermedia of different sorts this is beginning to change. Information space is undoubtedly becoming more of a place for entertainment, pleasure and socializing (Erickson, 1993).

On a general level, then, there are differences between cinematic and digital environments both in the possibility to interact, as well as in the way the spaces has come to be used by the society and culture. Although this fact has to be taken into consideration when comparing the two media, it does not prevent us from investigating more concrete levels such as visualization techniques. The purpose of this paper is to look into the techniques employed within the mainstream cinema, to see if they have any equivalents in computers and to speculate around whether these might be applicable in interface design. This will not be done in a general way, but specifically in relation to supporting navigation in information space (Benyon & Höök, 1997; Dahlbäck, Höök & Sjölinder, 1996; Darken & Sibert, 1997; Dieberger, 1994; Miller, 1993). How can film be instructive when supporting navigation in digital environments?

NAVIGATION AND ORIENTATION

Navigation is a mental and physical activity involving an environment (geographical, informational or social) and a navigator, where the navigator is attending to or monitoring the environment along some form of goal (more or less specified). With this goal in mind, whether it be a geographical place or a task, the navigator chooses a path or otherwise interact with the environment in order to reach this goal. All these activities might be assisted by some earlier conceptualization of the environment (cognitive maps etc.) or with some form of tool (maps, signs, other people). Sometimes the quality of the goal in mind is rather poor at the beginning of the navigation, i.e. abstract and vague, but is improved during the way. According to Downs & Steas's (1973) definition, wayfinding consists of orienting oneself in the environment, choosing the correct route, monitoring this route, and recognizing that the destination has been reached.

According to this sketchy definition the spectator of a mainstream film does very little navigating. Although she does travel through space in a transferred sense, she does not have any goal in mind (not even a vague one); she does not make route decisions; she does not physically interact with the environment. This is all done by the narrator. So why bother with navigation with cinema?

Well, cinema is extremely good at one dimension of the navigating experience, namely to *orient* the spectator. As we will see, the visualization conventions developed within mainstream fictive cinema aspires to make movement and spatial relations between objects and characters as clear and coherent as possible, and thereby prevent spatial ambiguities, disorientation and insecurity in the spectator (unless narratively motivated). According to this *transparent* style of film making, the spectator's *comprehension* of the story is the essential purpose of the visualization techniques, and the editing and other formal aspects is only means to convey this story, not ends in themselves. It is the story that should be visible to the spectator, not the visualization techniques themselves. The 'cinematic interface' (in its Hollywood version) is a *transparent window*, making the fictive world visible without attracting any attention (Andrew, 1984: 12 & 134).

The debate around cinema as an illusion and its strong potential to draw the spectator into the fictive universe (and thereby to some extent ideologically manipulate and affect her), is a sign of the success of the visualization techniques exploited by mainstream cinema. This is to be contrasted with the computer user whose attention to a very large degree is focused on determining the configuration of the system (and the like) rather than completing the specified task (Miller, 1993:10). The computer interface of today still applies a 'non-transparent style' (perhaps due to lack of standardization within the industry?).

Thus, orientation will in this paper refer to those visualization techniques that render the fictive space coherent, consistent and steady. These conventions seeks to disambiguate and stabilize relations and movements between objects, in order to present narrative information as effective as possible.

FILM HISTORY

However, in order to understand to what purpose these techniques have evolved, we must go back in film history, to the early cinema in the very beginning of this century (Gunning, 1994; Bordwell, Staiger & Thompson, 1985; Burch, 1990; Elsaesser, 1990).

For the first ten years after the first film exhibition in 1895, films were shown in vaudeville theaters, fairgrounds and in temporary premises, often by traveling projectionists. The films, ranging from actualities, nature films, panoramas, burlesque and entertainment acts, were only 1-3 minutes long - one reel length. A show was a potpourri of different sorts of genres, including musical accompaniment and often a lecturer describing and commenting on the contents. The style was very different from today. Cuts were seldom used and static cameras, almost always with distant framing, showed the whole of the scene and characters (fig. 1).

The camera respected the 'wholeness' of objects and actions and the space offscreen was unimportant. Spatial relations were established *within* the frame. This style was of course heavily influenced by the vaudeville-burlesque context. Most fictive films were vaudeville acts staged for the camera, exhibited in vaudeville theaters. The acting was rough and sweeping in order for the actor to reach those at the back of the theater in the long shot framing. But then again, this was perfectly natural in the vaudeville context.

Eventually there was a need for longer films and narratives, partly due to the industry's effort to erase early cinema's bad working-class reputation and make it a respectable middle class entertainment form. From 1905 to 1915 fiction films lasted around ten to twenty minutes, with increasingly more complicated, middle-class narratives. The camera crept closer, now beginning to use long shots, medium-shots and even close-ups in a regularly fashion. Closer views of faces and objects solved early cinema's problem with jumbled and messy long shots, where the viewer wouldn't really know what parts of the image contained the narratively important information. Cutting closer directed the spectators attention in a way that before was only possible with titles or lecturers. The close connection between early cinema and theater-vaudeville is probably one reason why it took ten years to leave the long shot tableaux technique (including its acting style). The complicated narratives also demanded more cuts - not only within a scene - but also between spaces and different temporal segments (think of such pivotal narrative techniques as *flashbacks*, *temporal ellipses* etc.). And segmenting the film also made it easier and cheaper to shoot and produce films, than would a single-shot film. Around 1915 full length films were produced on both sides of the Atlantic.

However, although closer-ups and cuts (of course in collaboration with acting and composition techniques) increased the articulation possibilities and hence the narrative potential, they also fragmented the long shot space and disoriented the spectator. Close-ups of objects make the offscreen space indeterminate and unclear, and cutting from one space-time fragment to another will jeopardize the temporal as well as the spatial relations between the two shots. How does the objects in shot 1 spatially relate to shot 2? Are they in the same room? Is this another scene? What is character X looking at? Is X (in shot 1) moving towards or away from Y (in shot 2)?

Some styles of film making liked and still are attracted to these disruptive effects and often play with them in order to disorient the spectator (avant-garde, art cinema, Russian cinema in the 30's etc.). American film industry had another aesthetic and commercial purpose: namely to tell narratives in the most renderable form possible. To them the disruptive effects of cutting and closer-ups had to be overcome (or 'tamed') in order to maintain a stable fictive space in which the spectator could concentrate on the narrative events. The formal aspects of the moving image should be *transparent*, in favor to the story. How was this accomplished? What techniques and conventions were developed within 'Classical Hollywood Cinema', in order to *support* the spectator (to use computer language) in this effort to understand the events depicted?

Some of these were concerned with directing the spectator's attention within the frame. These dealt mainly with lightning, composition within the frame (for instance depth staging) focus and camera techniques etc.. My focus here will instead lie on the orienting conventions involved in, on the one hand closer-ups and the offscreen space they generate, and on the other techniques to smoothen the disruptive force of the cut.

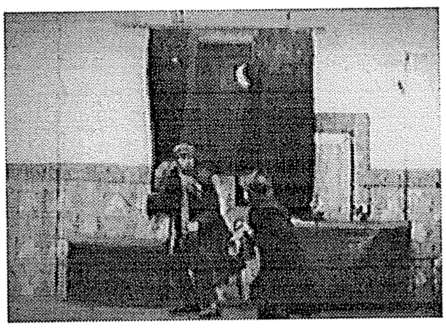


Fig. 1. *The Life of Charles Peace* (1902)

2. OFFSCREEN SPACE

Whatever its shape, the frame always makes the image finite and limited. From a continuous world, the frame selects a slice to show the audience, leaving out space to the left-right, above-below, behind the set and behind the camera (Burch, 1973). Offscreen space is that space not visible on the *screen*, but is still part of the *scene* (Bordwell & Thompson, 1993:495). However, this space, depending on the framing and cutting style, is more or less important. In early cinema, as we saw, the long shot tableaux style provided the spectator with the wholeness and totality of the scene. The limit of the frame was equivalent with the limit of the stage and almost everything taking place in the scene took place on the stage - and hence in the frame.

However, in the classical Hollywood style, the camera cuts closer and show the spectator tiny segments of the general scene: medium-shots and close-ups of faces, reactions, objects, parts of bodies etc.. This enhance the narrative articulation potential, but closer framings also present a orientation problem, since there are less spatial information in the frame. In order to prevent such effects the film often suggests or tries to indicate to the spectator objects and happenings *outside* the frame. These techniques seek to activate the viewer's imaginative inferences and her construction of (non-visible) objects and happenings in offscreen space.

a) Object continuity

In the real world we have a fair amount of knowledge of how objects and things look like. Adults expect objects to be whole and continuous even though only parts are present in our visual field. Although I only explicitly see the upper half of the chair in front of me (because the table blocks the view), I do not presume that the chair is cut in half and floats in the air, but that the chair continues behind the table, although I can not see it for the moment.

The same processes are present in comprehension of paintings, photographs and moving images. Although the frame ends at the neck of the man in fig. 2, the viewer does not interpret this as a picture of a decapitated head. She expects he has a body and that this continues outside the frame into offscreen space (like a window). She assumes that things in image space possess the continuity they usually have in real life. The 'naturalness' of conventions of cutting closer (in all types of visual arts), is probably due to our presumptions about a continuous space and knowledge about typical object appearances.

Other offscreen cues are not this obvious. Consider a hand coming into the frame, a shadow, or perhaps some cigarette smoke. They all signal the presence of someone in offscreen space, and they exploit the spectator's everyday conceptions about hands/bodies, shadows and cigarette smoke. These cues may trigger curiosity and other emotions: who is holding that cigarette? Who is the murderer?

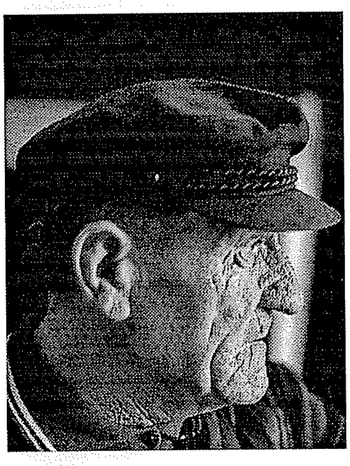


Fig. 2

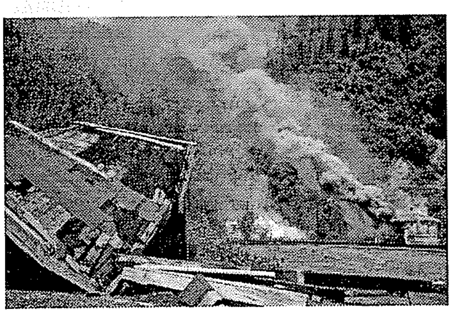


Fig. 3

b) Gaze direction - *Point-of-View editing*

In everyday life, following another person's gaze is a very basic and useful behavior (Butterworth, 1991). By looking at the same objects as the gazer is looking, you can guess where her attention is directed and perhaps from this speculate about her mental dispositions (beliefs, desires, feelings, intentions etc.). This behavior is not restricted to humans but show up in quite complex forms in animals as well, which indicates its fundamental function (Gómez, 1991; Ristau, 1991).

Gazing behavior in cinema and cartoons is extremely important in order to relate different shots and spaces, and activating offscreen space. Consider fig. 2 again. Following the gaze of the old man suggest to the spectator that he is attending to something. The viewer is not able to see it for the moment, but the image makes her assume the existence of some objects/event in offscreen space. The image activates a curiosity in the spectator. If we from this image cut or transfer to another image of, say, the burning house in fig. 3, then the viewer most likely infers that the man is looking at the house: when the shot of the old man is shown, the spectator assumes that there is something occupying offscreen space (right), and when the shot of house is shown, the spectator infers that the old man is somewhere offscreen (left?). This juxtaposition of spaces thus generates spectator inferences about the man's *point-of-view* and to what he is attending, but also that the two spaces are within sight's distance in the fictive space (that is, pretty close). This trope is extremely common in all genres of moving images.

It is interesting to note that 'the man looking at the house' is never explicitly present within the image frame. The spectator never *sees* the two together. The spatial relations between the shots only belongs to the imaginary, mental space that the spectator is *constructing* in her mind during viewing. In reality, these two spaces probably never met at all: the man may be shot in Stockholm, and the house in Edinburgh.

The spectator's inference concerning the spatial relations between the two shots, is thus not 'objective' in any sense of the word. It is an *interpretation* of the stimuli presented. But why does this interpretation seems so 'natural'? Why is this juxtaposition of images interpreted in the same way by most of us? For one thing, we have been exposed to this convention during many years of film watching. But the reason why the convention came in dominance in the first place, was probably due to the everyday gazing behavior competence in the spectator. Even before cinema, we were used to follow other people's gaze. The point-of-view editing convention exploited this knowledge in order to relate two fragments of space.

Gaze can connect spaces even more tightly. If a person looks offscreen and then the second shot depicts someone else also looking out the frame, then the spectator most likely infer that they *look at each other* (figs. 4 & 5). Here offscreen space becomes inhabited with another person. This type of editing (*shot-reverse-shot editing*) is one of the most commonly used editing devices in mainstream cinema, linking disparate spaces very effectively, and hence orient the spectator.

Fig. 4. *Notorious* (1946)

Fig. 5. *Notorious* (1946)

c) Sound

Research on the sound track within cinema studies has been marginalized in comparison to studies of the image track. The sound cannot be described and printed in books as easily as images, and many film scholars used to originate from art studies and their preferences lay more on the image side. Yet cinema was never silent. Music, singing, lecturers, live behind screen sounds etc. were from the very beginning pivotal elements in cinema exhibitions. In today's mainstream cinema, sound creates strong effects and yet remains quite unnoticeable by audiences. However, its relations to and reinforcements of the image track are many and complex and certainly worth studying (Bordwell & Thompson, 1993).

As far as cinematic offscreen space is concerned, it is hard to exaggerate the value of sound. Auditory information may indicate the presence of objects and events offscreen, and thus create expectations on the surrounding space. Consider again the shot of the man in fig. 2. Depending on the type of sound accompanying the shot, the offscreen space will contain different phenomena: a door being opened, someone talking, a barking dog or something burning. And moreover, the spectator will probably assume the man is attending to the source of the sound. Unless the filmmaker has some other purpose (e.g. create suspense as to the appearance of a monster or murderer), we expect to see to what the person is attending. The spectator anticipate an upcoming shot of something burning (e.g. a house).

Anyhow, in a style where the camera cuts closer every now and then, sound is extremely valuable to the filmmaker, not only for narrative effects, but also to orient the spectator by providing cues to adjacent offscreen happenings.

d) Spatial overlaps

One very common cinematic editing technique is *analytical* cutting. Here, the first shot of the scene gives a spatial and situational overview (*establishing shot* - fig. 6). Then the camera cuts or dissolves closer (to a medium/long shot in fig. 7, and then a medium shot in figs. 8-9) and finally ends with close-ups (figs. 4-5), often when the scene reaches its climax. Sometimes the scene is completed with a wide shot once again. This editing trope thus 'dissects' the space into fragments (hence the term *analytical*).

Fig. 6. *Notorious (1946)*

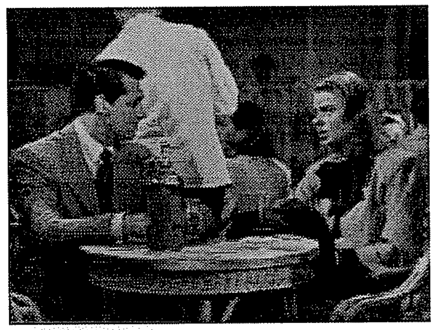


Fig. 7.

Fig. 8.

Fig. 9.

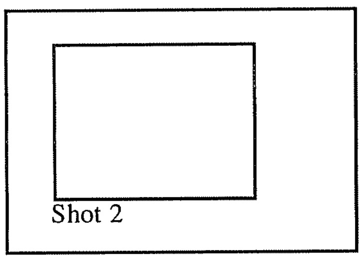


Fig. 10. Spatial overlap

The initial wider shot(s) undoubtedly have a number of functions. One of the most important, however, is to give the spatial setting for the closer-ups. That is, the establishing shots set expectations on the offscreen space of the tighter shots. This is called *spatial overlap* (fig. 10).

Consider the case of *Notorious*. If the filmmaker would have cut from a preceding scene to the close-up of fig. 4 the spectator would have had some 'orientational' problems: Where are we? With whom is he talking? What is the spatial relation between this shot and the former scene?

As the scene is edited now, we first get an overview of the street, then the table with *both* the characters. After that we cut closer to medium-shot, still with the other character's shoulder within the frame, until we reach the close-up of a face with no *explicit* clues to what goes on in the surrounding off-screen space (figs. 4-5). However since there have been spatial overlaps the spectator knows pretty well where the close-ups are set and how their offscreen spaces are structured. Spatial overlapping and overview shots provide information on offscreen space in closer-ups.

e) General expectations

Certainly, offscreen space is also indicated by appealing to general knowledge and expectations in the spectator. If the spectator knows a close-up is taking place in a restaurant (by the background sound, for instance), she can postulate things and objects in the surroundings prototypically connected to such a situation: tables, other guests, plates, forks, waiters, kitchen etc.. In order to come up with hypotheses of this sort, the spectator has to possess knowledge of a more general, everyday type, often culture specific and *script* like (Schank & Abelson, 1977).

Again, cinema exploits psychological dispositions and knowledge in the spectator in order to imply offscreen objects and happenings, and thereby contextualize the image and make it comprehensible and manageable to the spectator.

DIGITAL OFFSCREEN SPACE?

Is there an equivalent to off screen space in computerized environments?

Well, since there is an abundance of different types of digital environments, the question can not be answered generally. In many games (e.g. *Myst*, *Tomb Rider*, *Riven*, *Safecracker*), immersive environments, Cooperative Virtual Environments (*MASSIVE*, *DIVE*), multimedia productions and visual chat environments (*Palace*), the space is designed with reality and realism in mind (figs. 11-13). The graphics contain interiors or exteriors with everyday-like objects, such as rooms, carpets, beds, trees etc.. The user expects space and objects to have similar constancy and stability as everyday space of reality (with minor exception). In these cases, the frame limits the surroundings in the same manner as a camera does, and space is thought to continue beyond the frame. It is foremost here the cinematic techniques described above might have an application.

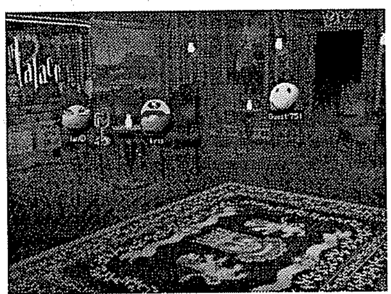


Fig. 11. *Palace*

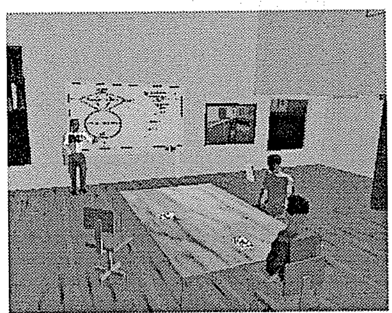


Fig. 12. *DIVE*

Fig. 13. *Safecracker*

In other cases the environment is more abstract and less space-like. Consider an ordinary interface of a personal computer (fig. 14). Here the background is abstract (neither interior nor exterior, no horizon or walls), the objects (buttons, menus, document icons, windows) do not look like or behave like everyday objects (they are flat, float around, make windows pop up), and things might happen here that have no equivalents in everyday life. In order to make your way through and understand this environment you have to have specially acquired knowledge - your everyday knowledge is not enough. And since space is abstract, the space offscreen (right-left or below-above) is not containing anything particular. The landscape or objects does not 'stretch out' into the distance in any direction. Everything of interest is contained within the frame.

But cannot the hidden information and 'spaces' behind menus, buttons, applications etc. count as an off screen space? The different icons on the desktop might be seen as indications of spaces offscreen ('help-space', 'configuration-space', 'preference-space', 'edit-space' etc.). Or in the same vein, what about a hypertext like the web page in fig. 15. Could not the links (in all their different colors) indicate the presence of various types of objects/information/events, off screen?

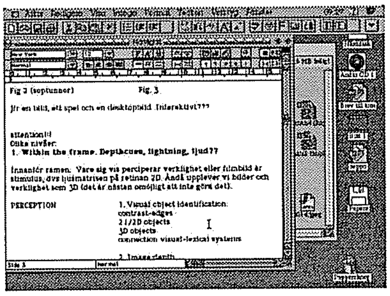


Fig. 14. Macintosh Interface

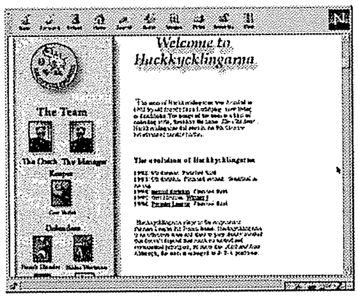


Fig. 15. Web site

The problem with calling this off screen *space*, is that the information, in contrast to cinematic offscreen objects, is not adjacent to the frame. When browsing the web, you do not conceptualize the information as being right outside the edge of the image. The information is not experienced as to the left-right or below-above the segment visible on your screen for the moment. The links, buttons or menus does not signal offscreen *space* - but rather offscreen *information*, and that is something much more abstract. This makes it more difficult to implement the cinematic conventions, since most of them depend on a 'realistic' space.

Although there surely exists middle course solutions between these two kinds of interfaces, they seem to be two categories employing very different visualization techniques. As far as offscreen space/information is concerned, they not only diverge in the level of abstractedness/realism, but also in the general experience. In cinematic and 'realist' environments the spectator/user is *drawn into* the world that seem to exists on the other side of that screen/interface. The user goes into the virtual world (through the 'window') and explore its hidden features. In contrast, in abstract interfaces the offscreen information is seemingly coming out *towards* the user: windows and menus *pops up*. The user does not go into menus or web pages - they seem to come out towards her. The direction of movement in showing the offscreen space/information appears to be different in the two cases.

With the distinction between realist and abstract interfaces in mind, let us see how offscreen information/space is indicated and used in different examples, and if the cinematic devises I talked about have any equivalents in interfaces. If we can better inform, or suggest to, the user what is outside the frame, will that support her experience of orientation?

a) Object continuity

Above I tried to show how closer framings limited disorientation in the spectator by exploiting everyday assumptions of space and object continuity. In abstract interfaces this device is difficult to put in work, since the objects and space does not obey 'the laws' of everyday space. For instance, although a word-processing document (*Word*, *Claris* etc.) or a web page in a browser aspire to live up to the 'piece of paper' metaphor, there are very few similarities between the behavior of a real and a virtual *document*. I have on several occasions observed fairly computer experienced users browsing my and other's web sites and was struck by their inability to understand that the information continued further down or up the 'page' (accessible through scrolling). They had little sense or expectations on offscreen information and considered the screen to display the whole situation (just like the early cinema spectator). Perhaps this inability is caused by the fact that the 'objects' on screen (web 'pages', 'documents' etc.) are not everyday objects and do not behave like such. Unlike the case of human bodies, bicycles and cars, the user does not have an conception of how a 'whole' web page or word document looks like (but see the hypertext documents in Páez, Bezerra da Silva-Fh & Marchionini, 1996). Would the problems be precluded if the design generated such a conception in the user?

Or take a menu, a button or a hypertext link. All of them hide some offscreen information, but there is no intuitive everyday visual cues to indicate what is hiding 'behind' them. They are all semantically determined with terms or symbols that has to be learned through repeated actions. There is no *natural design* here (Norman, 1988) and no object continuity, telling us what to do with these objects or to provide clues to the offscreen information behind them. (Then again, we have to ask ourselves if realistic space and objects are worth striving for. Is not the point of a word-processing document that it is *not* behaving like a real document, and that you could do things to it impossible to do in real life?).

We also saw how filmmakers used offscreen space to consciously hide important or frightening information (the face of a murderer/monster). By only showing single parts of an object, curiosity and excitement are generated in the spectator. Could this be used in a hypertext? Will a shadow of an elephant coming in from the right make the user curious of the contents of the adjacent page (and could she click on the shadow?)? Will cigarette smoke pouring in from below (accompanied with some sound) make the user wondering who dares to smoke in her computer environment (and what this person has to say)? And will this enhance the feeling of orientation? Will it support the user's navigational abilities, e.g. finding your way back to this node, memorize the path taken, make shortcuts?

b) Gaze directions

Because digital environments more and more are inhabited with agents, avatars and personal guides or assistants, I suspect gaze behavior to come in focus over the next years. In a world filled with anthropomorphizised creatures, gazing and looking will perhaps be the primary means with which to place psychology in them (*attending*, *perceiving*, *believing things*, *wanting things* etc.) just like gazing is used and understood in everyday life. Then we are getting close to the space of narratives and narrative games (e.g. *Tomb Rider* and other adventure games with characters). But will it support orientation, and if so, how?

There are a couple of cases with creatures in abstract space. *Microsoft's Office assistant* is one. This guide, which can come in many types of appearances, pops up once in a while and suggests solutions to problems you might be encountering for the moment. All communication is however done textually, except for different kinds of 'bodily' postures indicating surprise, sadness, anger etc.. But the direction of gaze in this assistant is never employed: it does never look at objects within or outside the frame/window.

In a more realistic example, *Palace* is a visual chat environment, with a realistic but static background image, which displays doors and openings on which the user can click in order for an new room to appear, including new chatters (fig. 11 above). Users are represented as colored blobs whose facial expressions could be manipulated by the user to indicate emotional reactions to what is said. Communication is however mainly done through the keyboard and displayed in cartoon balloons. Although the environment is not abstract, the avatars indeed are. The user can move them around, but turning your head/eyes towards someone or something is impossible. The complex and exciting gazing behavior that would take place in a real social space, is not present here.

In other cooperative environments like *DIVE* the avatars are detailed enough to exhibit gazing behavior. The user can turn around and change point-of-view (through keyboard and mouse commands), and other users can see in which direction she is looking. Here offscreen space and gazing direction cues work in similar manners as in cinema (and the real world).

Can the cinematic gazing convention support orientation? Well, to some users, only the presence of an humanoid creature in the abstract environment would have positive effects on the feeling of security and pleasantness (and thus to some extent orientation). But also, browsing through a hypertext with creatures of different types looking right or left at offscreen information space, would surely - for some users - be more fun (and possibly more efficient) than clicking on abstract links. If the user knew the creature and its preferences, or if the browsing was imbedded in a narrative (Persson, 1998), then the user could make inferences about the object of the gaze. Perhaps gazing would be complemented with semantic/verbal suggestions or with sounds from offscreen space?

Then again, for many users and situations, e.g. when you need to find a particular piece of information as fast as possible, such an environment would perhaps be more of an hinder than help. It is not my point to overthrow abstract interfaces, but only to suggest alternatives that might exists in parallel and perhaps primarily for pleasure.

c) Sound

Considering the enormous navigational and space creating potential of sounds, it is rather odd that it has not been exploited more in interface design. At best, the auditive channel is used for *foreground sounds* and synthetic warnings. But cinema exemplifies the possibilities of *contextual* and *background-ambient* sounds (Ferrington, 1996), particularly in relation to offscreen space. Sound does not only generate emotions, or change and enrich the interpretations of the explicitly visual, but most significantly creates expectations on the surrounding space. By designing a *soundscape* (Macaulay, Benyon, Crear, 1998) that made the user aware of offscreen phenomena such as the contents of adjacent nodes in a hypertext, or other users (visiting the same page as you are for the moment) etc., the orientation possibilities would increase for the user. Studying the techniques of cinema would in this context be a good start (Macaulay, Benyon, Crear, 1998). Although there is no place here for digressions, sound in cinema and interface design is definitely worth a closer study.

d) Spatial overlaps

Computer interfaces often use overlapping strategies. Different sorts of *index pages* (web sites, file systems, CD-ROMs etc.) basically have the same function as the cinematic establishing shot. An initial overview provide the user with some conception of what is to be found outside the frame of the subfile she will be visiting. The index page sets expectations on the offscreen information of the upcoming subpage.

This comparison does however halt slightly. Once again, the computer space is often much more abstract than cinematic space. The relations between nodes in an index page is less of a *spatial* character and more of a conceptual/structural one. In contrast to the scene from *Notorious* above, where the establishing shots provide the spatial surroundings and specify the offscreen space of the upcoming close-ups (in fig. 4 we know the woman is somewhere offscreen left), the index page does not establish *spatial* relations, but rather semantic/structural. When visiting a subpage the user cannot tell whether other subfiles is to the left or right, and this is typical of abstract space.

Could we support navigation if we spatialized the index- and sub-nodes, and if the movement from the former to the latter where experienced as a zooming-in or a cut-in? Many environment of 3D browsers seems to employ these techniques (e.g. eSCAPE's web environment in which every site is represented as a sphere and the navigator enters it and explores the contents on the inside). However, these spatial structures consists of abstract geometrical figures floating in space. The structures do not in any sense reflect the contents of the nodes. If spatialization is to help navigation (and the spatial overlap techniques from cinema be employed), I think this has to be done. There has to be some relation between the spatial environments and the content, just like there is a relation between the narrative and the objects - characters in the spatial layouts of cinema. How this could be applied in large and constantly changing information environments like the web, will be the critical challenge.

3. CINEMATIC EDITING TECHNIQUES

A cut means an instantaneous change from one framing to another. This includes a change of camera angle or the size of the space segment (long-shot/closer-ups). In either case the continuous space is disrupted and fragmented. Every cut means a new spatial position. In order to limit this disruptive force, mainstream cinema has established a number of editing techniques. All of these attempt to 'smoothen the cut' or make it less visible to the spectator. They strive to orient the spectator by indicating the spatial relations between the two shots.

a) Spatial overlaps

I have already mentioned some editing techniques in the former section. Cuts from an establishing shot to a smaller area of the scene is a very common way to smoothen the cut. This function is, I believe, a direct consequence of the fact that the wider shot sets expectations on the offscreen space of the tighter shot.

b) Gazing behavior

The same holds true for offscreen glances. Since looking beyond the frame sets expectations on

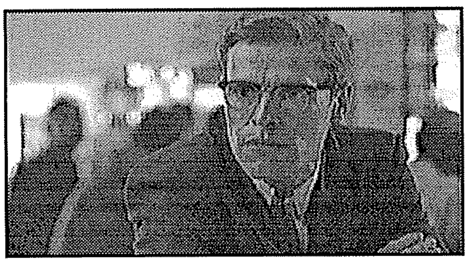


Fig. 16. *Miller's Crossing* (1989)

Fig. 17.

Fig. 18.

the offscreen space of the present as well as the upcoming shot, this smoothens the cut to a considerable degree. The spectator intuitively expects to see what the character is looking at and this makes the space-time disruption less observable and experienced. Point-of-view editing establishes the new scene as the object of the character's gaze, which spatially relate the two shots. In this way spatial disorientation is effectively inhibited. Gaze act as a 'mental glue' between two shots.

c) Movement

Match on action between two shots is also a very powerful device. This means that the cut is matched with some form of movement in the frame, for instance a character standing up, turning his head, opening a door or starting to run. In a sequence from *Miller's Crossing* the character is sitting at a table and turns around to watch the other people in the restaurant. There is a cut from a close-up (fig. 16) to a long shot from the opposite side or the room (fig. 18). The editor has many possibilities here. She can cut before the turning around, or she could wait until the movement is finished. In both of these cases the disruptive force of the cut will probably be most perceptible. Instead the editor has chosen to cut on action, with the character starting to move in the one shot (fig. 17), and finishing the turning in the other (fig. 18). In this case the cut comes in the middle of the action. The movement bridges the spatial break between the two shots, generates a visual 'flow' and constrains the disorientation effect in the spectator.

d) Sound

Audio atmosphere and sound is also a standard smoothening device in cinema editing. For instance, if the background noise of a restaurant is continuously heard over several disparate shots, the spectator assumes that they belong to the same restaurant space. Here the stability and continuity of the soundscape, compensate for the visual discontinuity between the shots.

Or if the sound of steps of a moving character continues from shot 1 to shot 2 (perhaps accompanying a match on action), it is easy for the spectator to be indulgent towards the disruptive image track. Although the image jumps and jerks, the soundtrack 'unruffle' these displacements, and makes it easy for the spectator to spatially (and temporally) orient herself in the new shot.

The importance of sound in these respects is forcefully demonstrated in many of the films of Jean-Luc Godard. Here the audio track is disruptive and constantly change 'position' (volume, intensity, soundscape), during a perfectly seamless visual channel. The spectatorial shock and disbelief towards the realism of space presented, indicates the power of sound in creating (de)stabilized space.

e) The 180° convention

In all of the above cases (except for sound) there is another underlying crucial convention, saying that directions within the scene form 'invisible' lines over which the camera is not to cross over. Gazes and movement are such directions. Consider the scene from *Notorious* above. In this scene the gaze line between the two characters forms the 180° line (fig. 19). An establishing shot (camera position 1) informs the spectator of the overall spatial arrangements: A is to left hand side of the image looking rightwards, and B is on the right-hand side of the image, looking leftwards. These directions of gazes are maintained in the closer-ups (camera positions 2 and 3), in spite of the fact that the angels of the camera have changed. However, if the camera would cross the 180° line to camera position 4, then suddenly B is looking *rightwards* (this shot is not included in the actual film - I have just inverted the image in camera position 3). This might create some form of disorientation: the shot seems to suggest that B is looking *away* from A and not towards him. Crossing the 180° line endanger the spatial coherence between shots and makes the spatial relations ambiguous.

Direction of movement can also constitute the 180° line. The editing in fig. 20 is a good example. If a character exists shot 1 going leftwards, the spectator expects him to enter shot 2 from right (still going leftwards) if spatial continuity is to be maintained. If the camera crosses the line however (camera position 3), then the character will enter shot 2 from the left hand side of the image suddenly going rightwards. Violating the 180° convention might in this case produce ambiguous interpretations and disorientation. Did the character change his mind and turn back? And if so, how did that happen? Or is he in another scene/place?

The 180° convention is however not rigid. There are in-between cases where crossing the line actually 'works' but then the crossing has to be shown explicitly (dolly shot, character moving over the line in one shot etc.). These rare cases might be valuable for anyone who seeks to explain the 180° convention, because although it is frequently used and described within production and cinema studies, no one has to my knowledge tried to give a psychological explanation of why this effect arises.

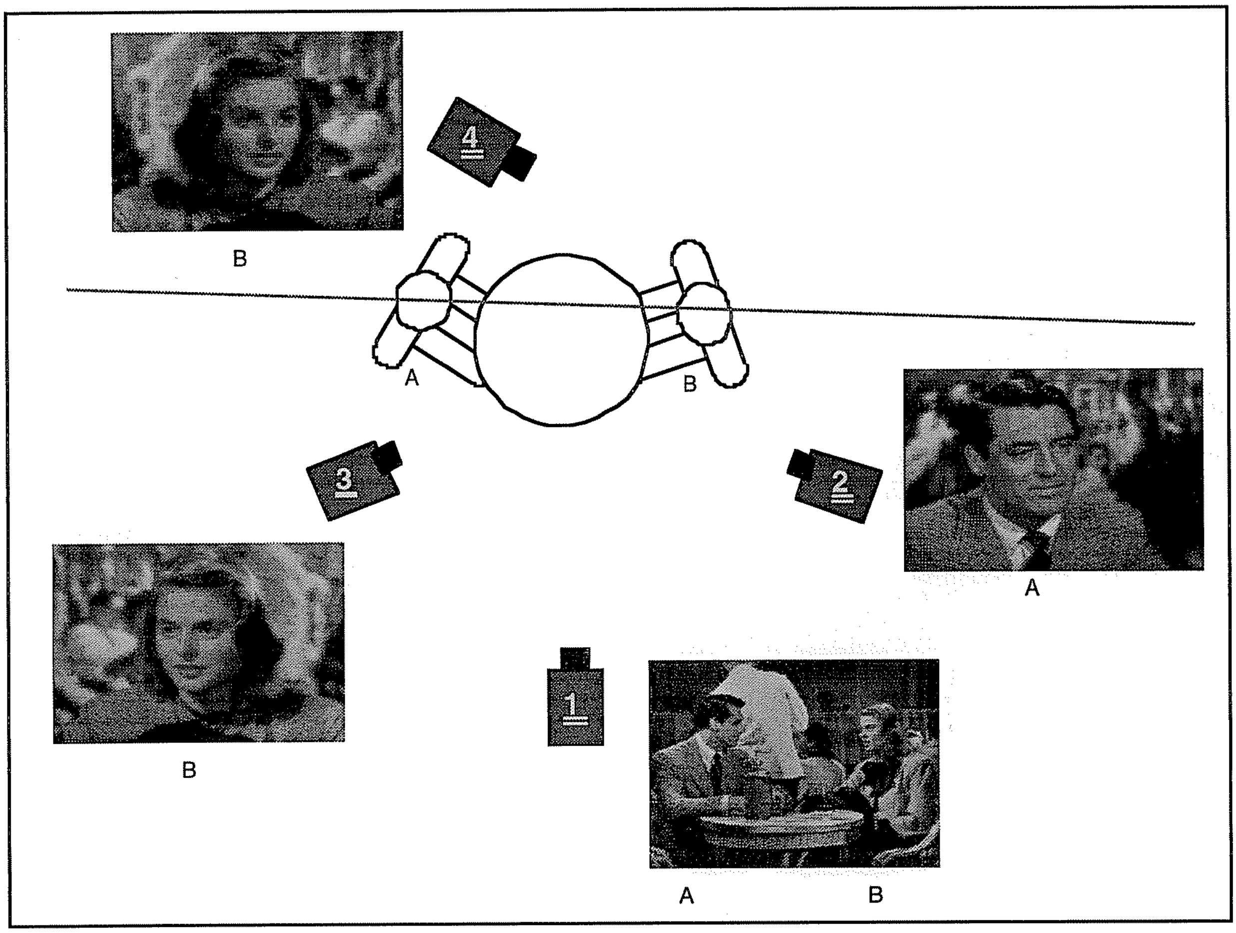


Fig. 19. 180° convention in *Notorious* (1946)

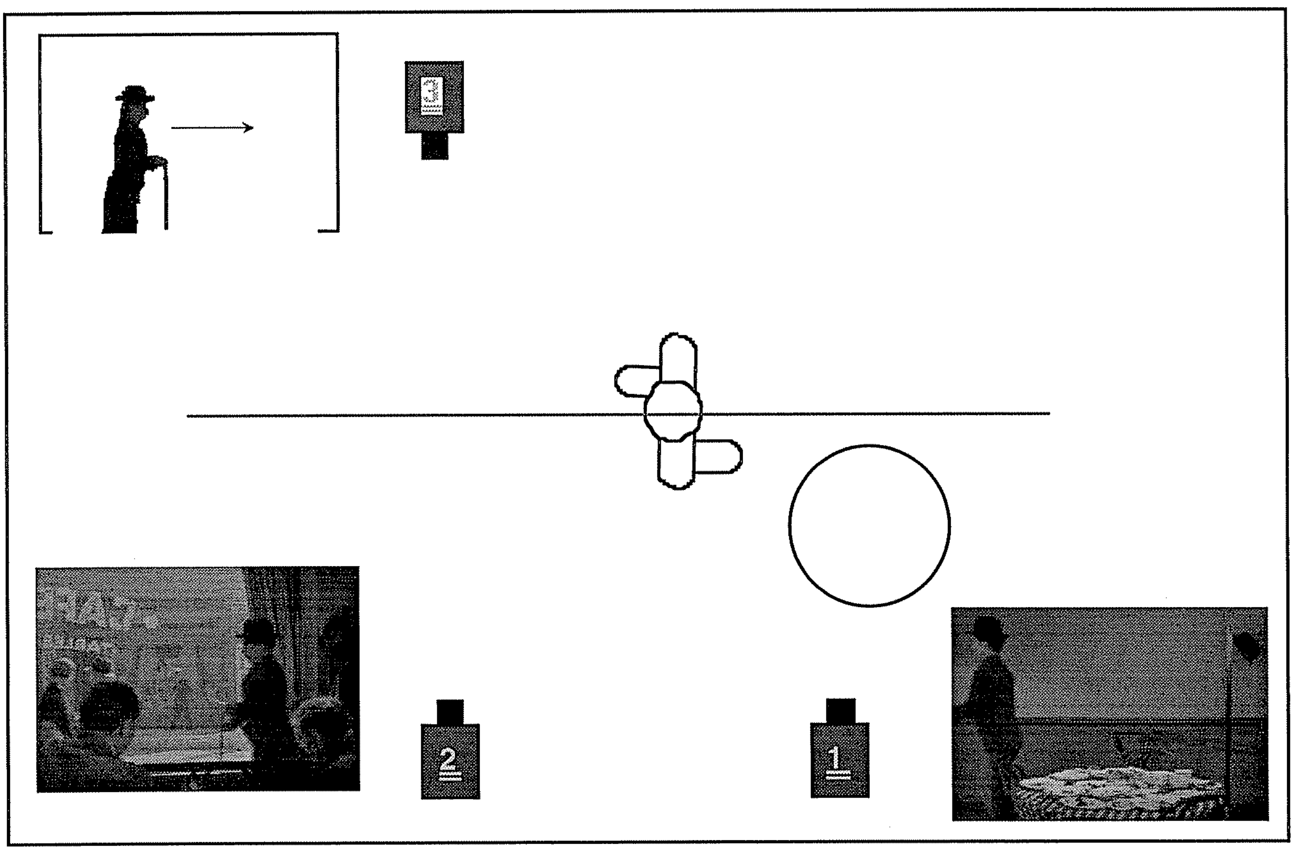


Fig. 20. 180° convention in *Modern Times* (1936)

The editing techniques considered here are forceful tools for the producer to link and relate sections of space and hence to constrain the disorienational effects of the cut. These techniques are in fact occasionally very effective. Many cinema/TV spectators are not even aware of cuts at all, but experience a spatio-temporal 'flow'. In spite of the very fragmented visual track, the space is experienced as coherent and disambiguous, thereby facilitating for the spectator to focus on actions and emotions of the characters and other important narrative events.

EDITING AND DIGITAL ENVIRONMENTS?

It is fascinating to image what computers could be like if interface design had adopted editing as a principle, just like film did 90 years ago. What would its potential be? It is hard to say.

But do not computer interfaces have their equivalents to cuts? What about a hypertext, in which the sites/nodes change abruptly and constantly present the user with new informational 'segments' and 'points-of-view'? And could not subnodes in a hypertext be seen as segments of that 'whole scene'? Or what about windows in a PC environment? When we change the front application window, is this not the same thing as cutting to another spatial framing?

Again, these environments are abstract enough to make it difficult to talk about 'space' and change framings within a scene. Instead these environments use other techniques to orient the spectator. The semantic information in the links and the user's voluntary motor activity (clicking) make the change smooth and expected. Color or other marking devices in the links also provide anticipations on what will come, and relate the information nodes. In a Mac or PC environment, the different *windows* are not views in different directions (left-right) like cinematic shots of a scene, but the windows are placed *on top of* each other (figs. 14-15). This makes the editing metaphor troublesome to embrace.

Admittedly, however, there are some similarities worth considering here, and I suppose, again, this is the point of departure of those developing 3D browsers and the like. What if nearby nodes were organized less according to content/semantics and abstract links, and more in a left-right manner? What if we laid out the information in an everyday landscape like a house or a city with familiar landmarks (and not in the abstract 'startrek style' most 3D browsers do), and made use of the many offscreen and editing techniques from cinema (e.g. 180° convention)? Would this reduce the feeling of navigational anxiety? Would such a solution overcame the lost-in-hyperspace problem? Again, I think that such a solution would possibly solve some navigational problems (orientation?), but far from all. Navigation and wayfinding is a rather complex phenomenon and we can not hope to fix these with graphics and visualization techniques only.

Or could we imagine a PC interface with windows that did not come out *towards* the user, but where the user *dived into* different spatial segment of the total 'working scene', and changing applications involved changing point of view? Would that sort out the messy PC interfaces of today? And how would the office assistant fit into this environment? These are all speculations and I have no definite answers.

In contrast to these abstract spaces, we have games, 3D and VR systems. Like cinema, this is (more or less) realistic space and we change view-points all the time. In some games and most VR systems the user's perspective is the camera and follows where the user go (either the 'camera' is placed in the very eyes of the avatar (*Riven*) with only hands visible, over the shoulder (*Tomb Rider*) or with the face of the avatar in full view). The perspective never leaves the user, and although there is a change of framings, this is a continuous movement and the user is in control of the camera all the time. The space is not fragmented like cinema's. The editing techniques described above are not needed.

In other systems it is possible for the camera to jump around. In games like *Full Throttle*, the editing is carried out with cinema style consciously in mind, and exploit many of the techniques I mentioned above. Since the player's control over the editing is minor (going into new rooms and environments) the general feeling is very much a cinematic one.

In some dynamic 3D environments used for scientific and educational application within molecular biology, computer engineering, medicine etc., the user is freed from camera control since the tasks performed are complex and demands total attention (Bares & Lester, 1997). In these systems the real-time camera planning is carried out by the system and entails selecting camera position and view directions in response to changes to objects in the environments, that are caused by the users' manipulations or a simulation. Since the system has to make on the fly decisions about camera angles and distances, it is perhaps in this context the editing techniques of cinema have their most useful application. In order to provide a continuous space and reduce the disorientation in the users, in order for her to execute the task as efficiently as possible, the smoothening devices of cinema would be worth considering (Bares & Lester, 1997).

4. CONCLUDING COMMENTS

Although I have only dealt with one set of rules here ('the Hollywood cinema'), this is of course not 'the right way' to make film. Filmmaker use conventions with different purposes in mind. 'Crossing-the-line effects' may be contra-productive if you strive for spatial coherence and orientation. This might even effect the understanding of what actually happens in the scene. If, on the other hand, the aesthetic or dramatic purpose of the film is to question the idea of a coherent space, to disorient or 'shake' the spectator, then breaking the 180° convention is most appropriate and a crucial tool. Conventions are like tools; they are in themselves neither right or wrong - they are *used.*

I think we can better understand what computer and digital environments are if we compare them with other artificial environments like cinema and fiction. It might even lead to some design suggestions that would not only make digital space more emotional and 'real-like', but in some contexts even support orientation (and possibly navigation). To some users in some situations these techniques will for sure just be a disturbance. In other situations they will have a value. Future work will hopefully better define what those situations and values are.

REFERENCES:

Andrew, Dudley (1984). *Concepts in Film Theory*, Oxford: Oxford University Press.

Bares, William H. & Lester, James C. (1997). "Cinematographic User Models for Automated Realtime Camera Control in Dynamic 3D Environments", *User Modeling '97*, Sardinia, Italy.

Benyon, D. & Höök, K. (1997). "Navigation in Information spaces: Supporting the Individual", Paper presented at *INTERACT '97*, Australia, July 1997.

Bordwell, David, Staiger, Janet & Thompson, Kristin (1985). *The Classical Hollywood Cinema: Film style and Mode of Production to 1960*, New York: Columbia University Press.

Bordwell, D. & Thompson, K. (1993). *Film Art. An Introduction*, New York: McGraw-Hill.

Burch, Noël (1973). *Theory of Film Practice*, London: Secker & Warburg.

Burch, Noël (1990). *Life to those Shadows*, London: BFI Publishing.

Butterworth, George (1991). "The Ontogeny and Phylogeny of Joint Visual Attention", in Whiten, Andrew (ed.), *Natural Theories of Mind. Evolution, Development and Simulation of Everyday Mindreading*, Oxford: Basil Blackwell, 223-232.

Dahlbäck, N., Höök, K. & Sjölinder, M. (1996). "Spatial Cognition in the Mind and in the World - the Case of Hypermedia Navigation", *The Eighteenth Annual Meeting of the Cognitive Science Society*, University of California, San Diego, July, 1996.

Darken, R.P. & Sibert, J.L. (1997) "Navigating large virtual spaces", *International Journal of Human Computer Interaction*.

Dieberger A. (1997). "Supporting Social Navigation on the World-Wide Web", *International Journal of Human-Computer Studies*, special issue on innovative applications of the Web, Vol. 46, pp. 805-825.

Downs, R. & Stea, D. (1973). "Cognitive Representations", in Downs, R. & Stea, D (eds.), *Image and Environment*, Chicago: Aldine (79-86).

Elsaesser, Thomas (ed.) (1990). *Early Cinema: Space - Frame - Narrative*, London: BFI Publishing.

Erickson, Thomas (1993). "From Interface to Interplace: The Spatial Environment as a medium for Interaction.", Proceedings of the European Conference on Spatial Information Theory, Heidelberg: Springer-Verlag.

Gómez, Juan Carlos (1991). "Visual Behavior as a Window for Reading the Mind of Others in Primates" in Whiten, Andrew (ed.), *Natural Theories of Mind. Evolution, Development and Simulation of Everyday Mindreading*, Oxford: Basil Blackwell, 195-207.

Gunning, Tom (1994). *D.W. Griffith and the Origins of American Narrative Film*, Urbana & Chicago: University of Illinois Press.

Macaulay, Catriona, Benyon, David & Crear, Alison (this volume). "Voices in the Forest: Sounds, Soundscapes and Interface Design".

Miller, Laura (1993). *Enhancing Usability in Virtual Environments: Understanding the Cognitive Skills Required in the use of Navigational Tools*, Dissertation, University of Central Florida.

Norman, Donald A. (1988). *The Design of Everyday Things*, New York: Currency Doubleday.

Páez, Linda, Bezerra da Silva-Fh., José & Marchionini, Gary (1996). "Disorientation in Electronic Environments: A Study of Hypertext and Continuous Zooming Interfaces", paper presented at *ASIS'96*, the annual meeting for the American Society for Information Science.

Persson, Per (this volume). "Supporting Navigation in Digital Environments: A Narrative Approach".

Plowman, Lydia (1996). "What's the Story? Narrative and the Comprehension of Educational Interactive Media", *Proceedings of ECCE8, the 8th European Conference on Cognitive Ergonomics*.

Ristau, Carolyn A. (1991). "Before Mindreading: Attention, Purposes and Deception in Birds?" in Whiten, Andrew (ed.), *Natural Theories of Mind. Evolution, Development and Simulation of Everyday Mindreading*, Oxford: Basil Blackwell, 209-222.

Schank, Roger C. & Abelson, Robert P. (1977). *Scripts, Plans, Goals, and Understanding*, Hillsdale, NJ: Erlbaum.

Chapter 12
Supporting Navigation in Digital Environments: A Narrative Approach

Per Persson
SICS

You do not see narrative and navigation in the same sentence every day. This theoretical paper tries to work out the connections between the two concepts and how studies on narratives can assist those on navigational issues in digital environments. I will argue that some dimensions of navigation and narrative overlap and that a narrative mode of organizing digital information environments, may be helpful and supportive for users disfavored by other modes of organization, e.g. spatial and semantic. Finally I will present some concrete design suggestions.

Supporting Navigation in Digital Environments: A Narrative Approach

Per Persson[1]
SICS Box 1263 S-164 Kista Sweden
per.persson@mail.film.su.se
http://film-server.film.su.se/Perhome/Index.html

ABSTRACT
You do not see *narrative* and *navigation* in the same sentence every day. This theoretical paper tries to work out the connections between the two concepts and how studies on narratives can assist those on navigational issues in digital environments. I will argue that some dimensions of navigation and narrative overlap and that a narrative mode of organizing digital information environments, may be helpful and supportive for users disfavored by other modes of organization, e.g. spatial and semantic. Finally I will present some concrete design suggestions.

INTRODUCTION: ORGANIZATIONAL MODES OF ENVIRONMENTS

Unlike real geographical space, digital environments can be manipulated and designed to an amazingly large degree. Depending on how the designer chooses to organize the environment, it will give rise to different types of experiences in the user/player/reader/navigator. Dourish & Chalmers (1994) identifies three major modes of information navigation used in present interfaces.

Some systems embrace the *spatial* paradigm and structure the information according to some geographical, 'real space' notion. For instance, in '3D' graphic CD-ROMs, visual muds and chat software or *Cooperative Virtual Environments* (CVE) and immersive environments, the objects are related spatially, like *left-right*, *upwards-downwards*, *below-above*, *outside-inside*, *in front of-behind* etc.. Here the navigator moves between the objects on the basis of their spatial connections and the experience is exclusively of a (geographically) spatial nature. These system seek to simulate real-world 'geographical' space in order to exploit spatial competence in the user, and thereby support navigation and orientation in the information environment.

Another 'space' parameter that is attracting interest is *social navigation*, where the movement from one item to another is cued by the activity of another or a group of users (see e.g. Benyon & Höök, 1997; Dieberger, 1997a & 1997b). If you could apprehend what other people in general or specific individuals were doing (e.g. what node they are visiting, if a node is crowded or not), this could influence your own goals and navigational strategies (cf. muds and CVEs). This experience would be less spatial and more like moving around in a public environment, attending to other people's behavior and intentions (e.g. crowding).

Probably the most common structure, however, is *semantic* organization, where the objects in the environment are related through some semantic connection like *similar*, *alike*, *more/less general*, *associated*. Here the navigator moves between objects on the basis of the semantic interpretation of the content in the nodes (the web, many CD-ROMs, software help systems etc.). The navigator, I would argue, does not primarily apprehend the connections between the nodes in spatial terms (*left*, *below*, *upwards* etc.) but semantically. This navigational experience is thus less spatial, but more like experiencing an encyclopedia or a book.

In most digital environments, Dourish & Chalmer's modes are of course more or less intermingled. What is important to emphasis here is, however, that different modes not only generates different experiences but also exploits, or are depending on, different abilities in the user. Spatial environments exploit spatial competence; semantic environments make use of semantic competence etc.. This is of course also to say that users with low spatial ability will be less successful in a system that exploits spatial competence. Hypermedia, data bases and hierarchical file systems have been found to be of a spatial character in this sense, disfavoring navigators with low spatial ability (Dahlbäck et al., 1996; Benyon & Murray, 1993; Vicente & Williges, 1988).

In this context *narrative* is one alternative organizational mode that may turn out to be beneficial to some users in some situations. It is the purpose of this paper to examine what a narrative organization might look like and how it might be supportive. Before we go into an

[1] Also affiliated with Department of Cinema Studies, University of Stockholm, Sweden.

analysis of narrative reception, however, we have to take a quick look at the processes involved in navigational activity.

NAVIGATION

Navigation is a mental and physical activity involving an environment (geographical, informational or social) and a navigator, where the navigator is attending to or monitoring the environment along some form of goal (more or less specified). With this goal in mind, whether it be a geographical place or a task, the navigator chooses a path or otherwise interact with the environment in order to reach this goal. All these activities might be assisted by some earlier conceptualization of the environment (cognitive maps etc.) or with some form of tool (maps, signs, other people). Sometimes the quality of the goal in mind is rather poor at the beginning of the navigation, i.e. abstract and vague, but is improved during the way.

It is not my purpose here to define the concept of navigation but the above will do for the moment. It fits well with Downs & Steas's (1973) definition of wayfinding: orienting oneself in the environment, choosing the correct route, monitoring this route, and recognizing that the destination has been reached. It harmonizes with the research on geographical world navigation within cognitive psychology, where the navigator apprehends the environment in terms of *landmarks*, *paths* and *regions*, and forms mental conceptualizations (or cognitive 'maps') of the environment in these terms. It broadens the perspective from only geographical cases to other environments (social and informational), which includes navigation in hypermedia, libraries, data bases and in society (although these spaces for sure implies different kinds of navigational activities - cf. Dahlbäck et al., 1996). It takes into consideration Benyon & Höök's (1997) distinction between *wayfinding* and *exploration* , where the former term refers to the situation where the navigator has a clear and quite precise goal or task, and the latter applies when the navigator only has vague notions of the goal or just want some general overview (or even just want to explore some objects) - cf. also Darken & Sibert's (1997) terminology. In order to be a navigational instance, however, there has to *some* form of goal on part of the navigator. Thus, evaluating a navigational tool not only includes how fast or effective the tool is in situations with clear and specified goals, but also how well the tool support the navigator in formulating a more qualitative goal ('what one really wants' - cf. for instance *Alta Vista* and *Yahoo* on this point).

The definition above also emphasizes navigation as an cognitively active process (monitoring, goal formulating, choosing route etc.), marking distance visavi the case of *transportation* where the subject knows the way extremely well or just follow some marked track or other people. Going by car to your new job might include navigation. Going there for the thousandth time does not. Neither if you go by taxi. The definition also excludes those instances when the subject travels in and monitoring some form of space but cannot interact with it or physically choose routes (cf. cinema or fictive space)

Navigation also often includes emotions of uncertainty or safety. If the environment or the tool does not give any positive feedback ('You're on the right track', 'Right now you are here and going in this direction') or if the navigator fail to attend to those, then the navigator will probably experience some form of anxiety and a feeling of being lost.

The navigational experience also includes some amount of *learning* the environment, and this will be the focus of this paper. By remembering landmarks/nodes and their interconnections, i.e. forming some conceptualization or model of the traversed environment, the navigator will be able to make a number of navigational activities: if coming to a dead end, the navigator does not have start all over again but only go back a couple of steps to a landmark and then take another path; take short cuts in the environment; perform better on next navigational session in the environment; be able to describe the way to others.

Generally then, my hypothesis is that the more connections between objects and landmarks a navigator constructs during a navigational session, the better she will accomplish these tasks. Making connections between landmarks will produce a richer mental representation of the environment and thus enhance memory (which will enhance navigation). This *mental network* will be a powerful mental tool in upcoming navigating situations.

Such connections can be triggered by different cues. They can be more or less *explicit* in the environment. Walking down a path may explicitly connect two landmarks in the navigators environmental model. A link between two nodes in a hypermedia may cue a connection in the mind of the navigator. And connections may have to be inferred on the basis of the position of other landmarks or nodes.

But, most importantly, the connections are of different types depending on the organizational mode of the environment. Spatial environments will cue spatial links and semantical environments will generate mainly semantic connections. Again, the type of environmental organization will demand different types of abilities and competence on part of the navigator, and some modes will fit some navigators (and situations) better than others. If the environment mainly seeks to trigger the mental construction of a spatial network, then navigators with low spatial ability may have problems in this construction work. It is in this context narrative organization becomes an interesting alternative.

NARRATIVE

The study of narratives has for the last eighty years been concerned with a wide variety of issues, but one important perspective is *reception studies*, dealing with the cognitive and emotional effects a narrative has on its reader/listener/spectator and how the reader actively applies different kinds of knowledge or cultural models on the text, draws inferences beyond the information given in the text and thereby construct some sort of coherence or understanding/interpretation of the information. This approach - let us call it *constructivism* - unites a wide range of scholars from cinema studies (Bordwell & Thompson, 1993; Bordwell, 1985; Smith, 1995), literature (Zwaan, 1993), communication studies (Messaris, 1994; Höijer, 1992) and psychology-linguistics (Graesser, Singer & Trabasso, 1994; Trabasso & Magliano, 1996; Bower & Cirilo, 1985; Mandler, 1984; Messaris, 1994; Bruner, 1990). In this approach narrative understanding arises in the mind of the reader in the interaction between some external structure 'out there' (text/discourse) and an interpretative instance (the reader/spectator/listener) with all its psychological and cultural etc. dispositions (e.g. schemas). The reader is actively searching for meaning and *coherence*, i.e. aspires to make a meaningful whole (*fabula* or *story*) out of a series of disparate events presented in the text (*syuzhet* or *plot*), assuming that there is some communicative *intention* behind the presentation of the events (from the author-narrator). Let's take a couple of examples.

" A woman is dreaming about cars. A window breaks. A profit is being made."

In this text the reader has difficulties relating the events presented. The causal, temporal and spatial connections between them are very vague and indeterminate. For most readers I would assume, there is no sense or coherence in the text as it stands. Consider, however another description of the same events:

"Sandra had dreamed about a new car for several months and last week she bought one down town. The next morning as she was going to work, the car wouldn't start . She got so angry she smashed the window of the car and her hand started to bleed. Next week she sold the car - with a profit."

Now we have a narrative coherence. The reader can relate the events *spatially*: Sandra buys the car down town; the window is probably being broken outside Sandra's house. Most importantly however, the reader is able to understand that the three events in the first text are a part of a series of *causes* and *effects*. Sandra's dream about buying a car is causing her to buy one; the broken car makes Sandra disappointed and angry which in its turn causes her to smash the window of the car; this causes her bleeding; and the whole event causes Sandra to reconsider her dream and sell the car. In this example the reader is also able to organize the events *temporally*: first she dreams; then she buys the car; the day after that she smashes the window; a week after that she sells the car.

Coherence and meaning are thus achieved when all/most elements in a text can be related causally, temporally and spatially. Psychologists like to see this coherence as some form of mental representation constructed by the reader - *mental text model*, or *situation model* (Graesser, Singer & Trabasso, 1994) - which includes text events and the explicit or inferred connections between them (like a network with nodes and links). The more established connections in the mental text model, the more 'tight' and coherent is the reader experiencing the text. If the connections, on the other hand, are loose or few, the narrative is experienced as loosely structured or even non coherent (as in many modern and post-modern narratives). Of course, there is no 'right' or 'objective' coherence for a given text, but this is all a matter of *interpretation*, that is, dependent on psychological, cultural, social etc. dispositions in the reader (which however might be similar for a great number of people; think of perceptual competence).

Of course the connections between the events might more or less explicit, demanding more or less inference work on part of the reader. For instance, the causal relation between smashing the window and the bleeding hand might be cued in many different ways (in order of degree of explicitness):

Smashing the window *caused* her hand to bleed.

The hand started to bleed *since* she smashed the window.
She smashed the window *and then* her hand started to bleed.
She smashed the window *and* the hand started to bleed.
She smashed the window. The hand started to bleed.
etc.

In the last examples, there is no causal relation *explicit* in the text, but this connection has to be inferred by the reader on the basis of knowledge about text markers ("and", "." etc.) as well as smashing windows and the danger it involves.

Of course, the strive for coherence and meaning is present in comprehension of all types of discourses, not only narratives (e.g. scientific papers, poems, recipes, manuals, news etc.). What is characteristic about narratives, however, is that the events prototypically involve *characters*. Narratives deal with antropomorphisized individuals with some form of inner psychology (they *think*, *feel*, *believe* and have *intentions*), who acts and reacts on the surrounding environment (which often includes other characters). The causal connections established in narrative comprehension are thus not 'scientific' or between objects, but rather of a folk-psychology type, with a back and forth movement between the inner dimension (or landscape (Bruner, 1986:14) of character psychology and an outer landscape of action, behavior and events (see Persson, 1997). The mental states of characters have effects on external action: Sandra's dreaming causes an intention which causes her buying the car; Sandra's rage causes her smashing the car. And outer events/action affect the inner states of the character: the fact that the car will not start causes Sandra's fury; the whole event causes Sandra to change her mind about cars.

Narratives also involve emotional effects on the reader/spectator. In fact, one of the main purposes of experiencing narrative - why we pay for books and cinema tickets - seems to be the affect or *pleasure* it produces. We expose ourselves to narratives because we want to be entertained (Brewer & Lichtenstein, 1982), excited, surprised, bewildered or frightened. We want to laugh, cry, trill, identify with characters and feel suspense, etc.. It is no coincidence that genre categories in both literature and cinema often are named after what emotional pleasure it seeks to achieve: *comedy*, *tragedy*, *thrillers*, *weepies*, *horror*. How these pleasurable aspects relate to the more cognitive aspects of narrative comprehension is still a unresearched area in narrative theory.

NARRATIVE AND NAVIGATION

Narrative surrounds us. In news, gossip, tabloid papers, cinema, commercials and literature, people are involved in narrative reasoning and comprehension. Some scholars even suggest that narrative thinking is a fundamental way to apprehend our everyday reality (Bruner, 1986 & 1990). Perhaps it is through narrative reasoning we 'come to grips' with the chaotic and changing reality.

Anyway, my main point here, and what is important in the following, is that narrative experience, through the constructive and pleasurable processes, *enhance experience of coherence and thereby memory and learning.*

Firstly, connections between events or units of information improves the memory of these events. A list of disparate pieces of information, without any relations (like the first text above), will be better learned if the reader could establish those connections. Networking, or *association*, is better than non-networking, as far as memory and learning are concerned.

Secondly, narratives provide alternative types of connections in this network. In fact, narrative construction processes involves a extremely rich set of inference processes, including causality, spatiality and temporality. In order to make sense of the information provided by the narrative text, the reader has to make use of not only spatial and semantic knowledge, but also everyday common sense causal competence (e.g. 'broken dreams and expectations can cause anger' and 'anger may cause violent behavior' - cf. from the text above). Since narratives typically deal with characters, the causal knowledge activated by narratives most importantly includes conceptions of how people's behavior causally relate to internal, psychological states (feelings, intentions, perceptions etc.) - cf. *folk* or *common sense psychology* (Wellman, 1990; Persson, 1997). If information space could trigger and activate these types of knowledge, instead of semantic or purely spatial ones, then this would enhance learning and remembrance for people with high narrative/social ability. *Narrative* networks are thus better than *non-narrative* networks, as far as those users are concerned.

An thirdly, if experiences are connected with affective states, they will generally be remembered better. Narratives are indeed such emotional triggers. Of course, making navigation more like a narrative experience involving emotions, suspense, humor, interest, surprise, curiosity, melancholia etc., is of course a reason in itself to look deeper into narratives. Information space will most likely turn more and more into a social and pleasurable space in order to attract users.

Narratives have and will have a central position in these efforts. However, narratives will also enhance memory of the specific navigational session, since they involve emotional experiences. As far as memory is concerned, emotions are better than non-emotions.

All of these features of narrative enhance the capacity for learning an information environment, and could be an alternative for those with low spatial or semantical ability. Learning and remembering are of course important aspects of pedagogical and educational research, and there has been great efforts in exploring how a 'narrativization' or information might make learning more fun and stimulating, and thereby better (MENO, 1996; Risden, 1997). But what is the relation between enhanced learning and navigation?

First, like I said before, if a navigator learns fast and remembers well the organization of an information structure, she will be better equipped to deal with dead ends and make shortcuts. Also upcoming navigational sessions will be handled more effectively than the first one, and the experience of *security* during the navigation (orienting oneself in the environment, and monitoring one's position) will probably enhance. Learning and remembering an environment are prerequisites for the navigator's familiarity with this environment. If a narrative organizational mode enhances memory and fastens learning of a given information structure, then this will probably also effect the navigational performance.

Secondly, if the navigator remembers the browsing session, she will be able to tell others about it later. Social navigation of this sort, where a navigator uses other people and their presumed knowledge as a navigational tool or filter, seems to be more common than has been thought of before (see cf. Dieberger, 1997a & 1997b). If narrative form enhances memory, then it will also facilitate for those processes taking place between browsing sessions and between people. Not only will the narrative mode make the navigator remember better, but it might also be more *describable* than other semantic-abstract modes. Everybody knows words for contents of a story, but not all will be familiar with terms for units in a semantic or abstract interface (think of *menus*, *index page*, *node* etc.).

Narrative seems to be a natural mode for interpersonal communication of this sort. Even in purely spatial descriptions, the narrative mode is sometimes used in order to facilitate memory. Höök (1991) collected and studied human route descriptions of Stockholm, directed at tourists with no experience of the environment. On some occasions the subjects added a little story to an important building or landmark, in order to facilitate memory of that landmark. Narrative seems thus to be a natural mode which people use in order to better memorize spatial descriptions, and if the system would support such narrative associations, many people would perhaps enhance their ability to memorize and generate descriptions as well as memorize them for own navigational purposes.

Making information space into a narrative space will thus, speculatively, generate a more emotional experience. Narratives will also in different respects enhance learning and memory for some people in some situations. This could support these people when navigating through what we today call 'information space'. But how is this to be implemented concretely? How can system developers trigger users to make narrative connections between information units (hypertext nodes or whatever)? How to make the browsing or navigational session less like a *traversal* and more like a *progression* or *development*, including narrative connections of different sorts?

First of all I think that the narratives have to be connected to the semantics of the information and not only to the formal structure. In a hypertext, for instance, there is not much point in having stories 'floating above' the information in the nodes, disconnected from what is said/shown in the nodes. It will be the narratives in association with the semantics that provide the enhanced memory effect. Although it does not have to be very tight and perfect, the mental connection or association between story and hypertext information has to be there.

Secondly, I believe that the narratives told has to be the same for everyone who comes to a particular node, and for every time. To change the structure in adaptive systems between visits has proven to generate confusion and feeling of being lost (Höök & Svensson, 1998). Carter (1996), for instance, pre-structured a hyperspace in several different ways, all reflecting the domain content. During the experiment, the system would switch structure when the subjects turned to a new question that would be more easily solved with the other structure. Users disliked the system and performed worse, possibly because they were unable to find their way back to landmarks and other navigational units that had been established and learned during the former sessions.

The analogy to narratives in information space is clear. If stories are to be used as 'landmarks' in a narrativized information space, they have to be rather stable and unchanging in order for the navigator to recognize them and orient oneself. If you suddenly change the appearance of these landmarks without warning, it will probably produce confusion and lessen the experience of security.

Thirdly, introducing narratives into an interface may seem like a laborious design project. There is, however, reason to believe that triggering narrative construction processes in the user may be accomplished by very simple means, graphically or whatever. Heider & Simmel (1944) showed people an animated film with abstract black figures, like triangles and squares, moving about on a white screen. Afterwards, the subjects were asked to describe the happenings of the film. Although not encouraged use narrative terms, this was often included into the descriptions. In order to comprehend the abstract course of events, the subjects gave it a narrative form (the big triangle 'chased' the small one because the big one 'wanted to punish' the small one for what the small one had done to the square, etc.). Among other things, this experiment shows that people are eager to generate narrative interpretations of the environment, even though the 'interface' is rather poor and abstract. It does not take much to trigger imaginative processes in the user. *The Tamaguchi*, combining an enormous success and a VERY simple interface, is another example of this fact.

CONCRETE DESIGN SUGGESTIONS

These are design principles, but before we come to an end I would like to sketch some concrete design ideas (which have been developed in cooperation with prof. David Benyon).

1. In a hypertext system you click on a link. A little dog (elephant, cat, vagabond, individual) appears, looking puzzled/happy. 'Oh right, you are going to check out what to find in Stockholm, are you? Be seeing ya!' ....and zooms off the screen.

You probably see this little dog quite often and over time get to know its character, facial expression, etc. He might pop up anywhere (rather unexpected and nonconsistently) and comment on the information in the destination node (on the basis of key word analysis). He might even suggest a story that in some vague sense is related to the site you are at ('Stockholm! I was in Stockholm once and I met the strangest man... Do you want me to tell you?' or 'I visited the Hedvig Eleonora church there three years ago! Are you able to find a picture of it somewhere?')

This storyteller would be equipped with a host of story lexicon, possibly with combinational possibilities (perhaps in contact with an on-line story-database, where stories are provided by users (see for instance <http://www.bubbe.com/> or <http://www.cyberenet.net/~sjohnson/stories/>). It would be possible to switch him/her on and off. I imagine the narrator to be visible in a particular space in the browser. Does it tell the stories in natural language or in images? Or both?

Of course, the user can be provided with different sort of background information of the storyteller. There could also be many different storytellers, each of them liking one genre of stories (comedies, tragic stories, action etc.). Thus the user can choose according to her genre preferences.

There would also have to be some sort of logging of what stories are told to a specific navigator, in order to avoid unnecessary repetition.

2. Another idea would be to have an ongoing narrative during the whole navigation session. For instance, on each page visited the narrator would present some progression or development of the story that in some sense was related to the information at that page. To continue our example from above, when coming to a site dealing with Stockholm sights, the narrator could (again on the basis of key word analysis) set the happenings of the narrative in a Stockholm environment, or introduce a character originating in Stockholm.

If the navigator want to go back to an earlier page, she may use the story and its progression backwards to find the right page. From there she might take another path, with a different angle to the story.

Eventually the narrative will come to an end and then the navigator has to choose to close the navigation session or to start another story.

How the story should be presented is an open question. Perhaps the narrator is telling you verbally, or maybe the narrator is absent and the story is told with moving or static images, or in a comic strip format.

This idea does not harmonize with the above principle concerning the stability of the structure: if you return to the same page on another occasion, the story told will most probably be of another one. Such an idea will not support social navigation.

3. Another, less ambitious idea, would be to create some form of narrator that informs the navigator what to find on adjacent nodes. The computer could check out all the links from a page whilst the user is reading the page, use a keyword count to measure the semantic content of those other pages and form that into a simple narrative. If it has knowledge of the individual user so much the better. You could imagine a system which counts occurrences of keywords, orders them by frequency and then uses a standard narrative structure to say something like..

> well if you go to <link> you'll find lots of stuff about <keywords> whereas
> <link> seems more concerned with <keywords>...

This could be combined with the *ALEXA* system idea: tracing other peoples' movements from the site you presently are on: 'Lots of people seemed to find Umeå University interesting when they visited the SICS site. Would you like to try?'

CONCLUSIONS

For many navigational and information seeking situations, a narrative organizational mode will for sure just be a disturbance. Users with high spatial and semantic ability the ordinary modes will be perfectly sufficient and probably a lot *more* efficient. When the goal of the navigator is clear and well specified, for instance a unique piece of information, the narratives in the information space will just be in the way. Narratives will not affect or support central navigational aspects (for instance how to formulate your goal, choosing the correct route, recognizing that your destination has been reached etc.).

However, when the user is spatially poor or when the goal of the navigation is just to explore a region of information space, then narratives may not only help, but also evoke pleasures, satisfaction and feelings of security. Users who perform badly in semantic or spatial modes of information organization, may perhaps use the narratives when identifying landmarks, taking shortcuts or going backwards.

Narrative modes of organizing environments will surely not be a substitute, but an additional complement.

REFERENCES

Benyon, D. & Höök, K. (1997). "Navigation in Information spaces: supporting the individual", Paper presented at *INTERACT '97*, Australia, July 1997.

Benyon, D. & Murray, D.M. (1993). "Adaptive Systems; from Intelligent Tutoring to Autonomous Agents", *Knowledge-based Systems*, 6 (3).

Bordwell, David (1985). *Narration in the Fiction Film*, London: Routledge.

Bordwell, D. & Thompson, K. (1993). *Film Art. An Introduction*, New York: McGraw-Hill.

Bruner, Jerome (1986). *Actual Minds, Possible Worlds*, Cambridge: Harvard University Press.

Bruner, Jerome (1990). *Acts of Meaning*, Cambridge: Harvard University Press.

Dahlbäck, N., Höök, K. & Sjölinder, M. (1996). "Spatial Cognition in the Mind and in the World - the Case of Hypermedia Navigation", *The Eighteenth Annual Meeting of the Cognitive Science Society*, University of California, San Diego, July, 1996.

Bower, Gordon H. & Cirilo, Randolph K. (1985). "Cognitive Psychology and Text Processing", in van Dijk, Teun A. (ed.), *Handbook of Discourse Analysis (vol. 1)*, London: Academic Press, 71-105.

Brewer, William F. & Lichtenstein, Edward H. (1982). "Stories are to Entertain: A Structural-Affect Theory of Stories", *Journal of Pragmatics* 6, 473-486.

Carter, Edward (1996). *Quantitative Evaluation of Hypertext: Generation and Organisation Techniques*, Ph.D. thesis, University of Edinburgh.

Darken, R.P. & Sibert, J.L. (1997) "Navigating large virtual spaces", *International Journal of Human Computer Interaction*.

Dieberger A. (1997a). "Supporting Social Navigation on the World-Wide Web", *International Journal of Human-Computer Studies*, special issue on innovative applications of the Web, Vol. 46, pp. 805-825.

Dieberger, A (1997b). "Navigation metaphors and social navigation in information spaces", Position paper for the CHI'97 workshop on Navigation in Information Spaces. (Available at <http://www.lcc.gatech.edu/~dieberger/Publications.html>).

Dourish, P. & Chalmers, M. (1994). "Running out of Space: Models of Information Navigation", paper presented at HCI'94 (British Computer Society).

Downs, R. & Stea, D. (1973). "Cognitive Representations", in Downs, R. & Stea, D (eds.), *Image and Environment*, Chicago: Aldine (79-86).

Graesser, Arthur C., Singer, Murray & Trabasso, Tom (1994). "Constructing Inferences During Narrative Text Comprehension", *Psychological Review*, Vol. 101, No. 3, 371-95.

Heider, Fritz & Simmel, Marianne (1944), "An Experimental Study of Apparent Behavior", *American Journal of Psychology*, 57 (2), April, 243-59. Republished in Aronson, Elliot & Pratkanis, Anthony R. (eds.) (1993). *Social Psychology, vol. 1*, The International Library of Critical Writings in Psychology, 279-295.

Höijer, Birgitta (1992). "Reception of television narration as a socio-cognitive process: A schema-theoretical outline", *Poetics* 21, 283-304.

Höök, Kristina (1991). *An Approach to a Route Guidance Interface*, Licentiate Thesis, Department of Computer and Systems Sciences, Stockholm University.

Höök, Kristina & Svensson, Martin (1998). "Evaluating Adaptive Navigation Support", unpublished paper of PERSONA project, SICS, Sweden.

Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*, Chicago: The University of Chicago Press.

Mandler, Jean (1984). *Stories, Scripts, and Scenes: Aspects of Schema Theory*, Hillsdale: Lawrence Erlbaum Associates, Publishers.

MENO (1996). For information see <http://www-iet.open.ac.uk/iet/MENO/meno-home.html>.

Messaris, Paul (1994). *Visual Literacy. Image, Mind, & Reality*, Oxford: Westview Press.

Persson, Per (1997). "Making Sense of Characters: Spectators and Folk Psychology", Dissertational chapter, Department of Cinema Studies, Stockholm University.

Risden, Kirsten (1997)."Can theories of traditional narrative understanding inform the design of hypermedia stories?", paper presented at MENO's Workshop on Narrative and Hypermedia. Accessible at <http://www-iet.open.ac.uk/iet/meno/HT97/kirsten.html>.

Smith, Murray (1995). *Engaging Characters. Fiction, Emotion, and the Cinema*, Oxford: Clarendon Press.

Trabasso, T. & Magliano, J.P. (1996). "Conscious Understanding During Comprehension", *Discourse Processes*, 21, 255-287.

Vicente, K.J. & Williges, R.C. (1988). "Accommodating individual differences in searching a hierarchical file system", *International Journal Man-Machine Studies*, 29, 647-668.

Wellman, Henry M. (1990). *The Child's Theory of Mind*, Cambridge: The MIT Press.

Zwaan, R. (1993). *Aspects of Literary Comprehension. A Cognitive Approach*, Amsterdam: John Benjamins Publishing Company.

Chapter 13

Navigation In Graphical User Interfaces

Rod McCall
Napier University

Graphical user interfaces can be thought of as information cities (Dieberger and Tromp). In such an environment users may seek to complete a given task in which they are in essence travellers moving towards a given destination. This paper will focus on the spatial aspects of GUI's and how support for navigation already exists or can be built in. The aim being to support as quickly as possible the users move from route to survey knowledge and wayfinding to transportation. To support the move emphasis is placed on task based interfaces which support the creation of spatially based mental models. In order to support task based interfaces emphasis is also placed on contextually based interfaces which support transparency.

Navigation In Graphical User Interfaces

Rod McCall

Dept. of Computing, Napier University, 219 Colinton Road, Edinburgh, EH14 1DT, UK

rmccall@dcs.napier.ac.uk

http://www.dcs.napier.ac.uk/~rmccall

Graphical user interfaces can be thought of as information cities (Dieberger and Tromp). In such an environment users may seek to complete a given task in which they are in essence travellers moving towards a given destination. This paper will focus on the spatial aspects of GUI's and how support for navigation already exists or can be built in. The aim being to support as quickly as possible the users move from route to survey knowledge and wayfinding to transportation. To support the move emphasis is placed on task based interfaces which support the creation of spatially based mental models. In order to support task based interfaces emphasis is also placed on contextually based interfaces which support transparency.

Introduction

In the real world we seek to find our way around in an environment. The environment can be within a building a city, the narrative in a film (Persson 1998)or even finding a document on a desk.

This paper will summarise some aspects of navigating within computer based environments in this case graphical user interfaces. The purpose of this is to focus on GUI's and see how aspects of navigation from real world and other computer environments such as hypermedia, MUD's MOO's and virtual reality etc. can provide theories which are relevant to GUI's. Further to this the concepts outlined here may be included in a navigational instrument, which it is hoped will provide a method for designers of GUI's and other computer based environments to build-in navigation and also assess it's existence within an interface system.

Background

Related Aspects of Cognitive Maps, Hypertext and City Form.

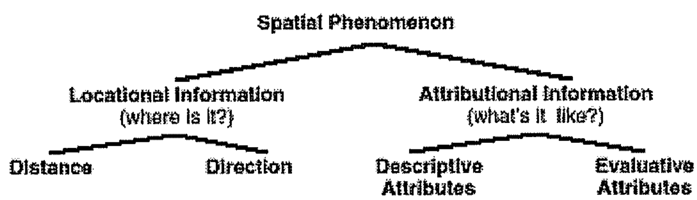


Figure 1:The contents of a cognitive map containing two types of spatial information (locational & attributional) from Shum.

The cognitive map in figure 1 illustrates some of the properties which may be required to navigate within an information (or hypermedia) environment. A spatial phenomenon in the example given could be a building or a sign. These can also be made broadly applicable to GUI type environments. For example distance and direction can both be applied to using the Wizards found in many Microsoft products i.e.

Distance: the number of dialogue boxes which I have to travel through in order to get to the one which contains the action I require to change from the default.

Direction: Direction already exists to some extent within GUI's. It can be found to a limited degree in the use of previous and next buttons within Wizards. The next button gives the idea that the user is travelling towards an intended destination, conversely back gives the idea of re-tracing previous steps. In addition to this some Wizards or dialogue structures have selection mechanisms which alter the ultimate path or direction through the dialogue structure. For example selecting a specific option may allow the user to enter a section of sub-dialogues on a specific option.

Descriptive attributes: a button of 40*50 pixels consisting of red text and right facing arrows.

Evaluative Attributes: By selecting an item from a dialogue using a series of check boxes the user is making an evaluative decision based on their own personal preferences.

Spatial navigation not only requires the consideration of physical attributes (distance, direction and descriptive) but also the context in which the user is working and also their preferences (evaluative). Hypermedia systems such as Sepia (Streitz, Haake et al. 1992) illustrate some of the evaluative aspects which can impact the decisions taken by users. The Sepia system is largely based on the concept of information triage where social and temporal aspects play a part in constructing and navigating within information environments. In essence therefore the Sepia system illustrates some of the properties found in the "information city" (Dieberger and Tromp). Where people are carrying out tasks in an ever changing and information rich environment. The precise way they carry out a task is dependant on the dynamic and environmental context in which they are placed. This follows in to the ideas advanced by Lynch (Lynch 1982) of layout of buildings within cities. The concepts advanced by Lynch can in part be seen as the physical aspects of the environment, for example landmarks, districts, paths and edges. Where landmarks are visible from a distance thus allowing the traveller to locate a desired area visually. The landmark may also lead to a district, an area which contains related buildings which have similar purposes are grouped. The interaction and activities which occur in a city are due to the evaluative attributes (Shum 1990) which are implicit in the physical form of the city. This followings directly into the ideas being advanced by the information city (Dieberger and Tromp 1996) where interaction within the environments is dependent upon the social context and preferences of the traveller. Hence while the work of all three (Shum 1990) (Dieberger 1996) (Streitz, Haake et al. 1992) remain distinct and separate in many respects there are several similar aspects. In essence therefore the ideas proposed by Shum (figure 1) can be thought of as integrating in with the ideas of Lynch. The Lynch concepts are the wider principles (districts , boundaries etc.) and those advocated by Shum, direction, description etc. are the more finite aspects. For example the traveller in the city is aware of the districts and boundaries but constructs a map in order to move within that environment. The map may consist of a spatial phenomenon as an example a church with a tall spire which acts as a landmark.

Landmarks and Districts in Graphical User Interfaces.

Navigating in computer environments (MUDS's VR, GUI's and hypermedia) requires the use of spatial knowledge and awareness. The required spatial knowledge is not unlike that required for non-synthetic environments. Spatial awareness is a key property of everyday life as people seek to navigate in towns and cities. Within cities people seek to use various cues as a means of finding their way around. In the urban planning context (Lynch 1982) here exists a number of elements which all aid in navigation such as districts, landmarks, nodes, paths and edges (boundaries). Navigational properties such as these can be seen to arise in graphical user interfaces, and can be found in common applications such as Microsoft Word. An example of districts exists in the toolbar layout which groups related icons together such as print, open and save. Further to this paths can exist, for example when a user resorts to using Wizard in an application. The previous and next buttons as well as the other widgets used for options provide direction and create a path through the given task. Further to this the precise path taken through the dialogues may change as different options are selected. Landmarks are also an important feature of graphical user interfaces, for example the default GUI layout in Word® offers several landmarks where users can approximate what they are likely to find in a given interface location. For example the interface offers several sub-menus (or districts), where each menu holds related options. Further to this a menu may also be seen as a landmark due to the fact it can be used by the user to locate a given district. Further to this using a Wizard (as found in many Microsoft applications) as an example, a landmark may also exist in the form a series of options which the user regards as critical. The critical options may provide the user with some sense of the level of completion within the Wizard, thus acting as a landmark or goal which they move towards. It is not being suggested that Word is the ideal interface, but rather it does exhibit some aspects which may be regarded as navigational cues.

The Role of Language in Navigation

A graphical user interface is an information rich environment, somewhat similar to the information city idea as proposed by Dieberger (Dieberger and Tromp). In another paper on enhancing the ability to navigate on the web by using text based VR, Dieberger (Dieberger 1996) also illustrated the value of using a rich language to aid in communication. The use of a rich communication language permits improved interactions between tools, agents and perhaps other users, therefore introducing a level of social navigation (Svensson 1998). An example of social navigation in a real world environment is when a traveller asks a native language speaker for directions. Frequently this interaction will take place using only the name of the area of interest, then the local person will have to attempt to feed back the instructions using the native language, where large chunks of the conversation may be simplified. The result is often a slower process of communication which can also be quite confusing for both parties. Therefore in order to permit social navigation both with other users, the interface in general and agents a language must be used which permits easy and effective communication.

Transparency and Contextual Awareness with the Interface

Interfaces which use grouping of menus and buttons are in essence imposing a map upon the user. In doing so these interfaces are attempting to encourage the user to think in terms of an overview (survey knowledge) of the interface. Extending this further the navigational cues must in essence move away from supporting purely atomic interaction (Gentner and Neilsen 1996) which is a by-product of direct manipulation interfaces. The reason for moving away from direct manipulation is that by forcing users to stick with an atomic style of interaction, they are frequently dividing attention away from the actual activity to dealing with the interface. In doing so attention being divided among the activity (e.g. changing the style of a document), conceptual interface view (i.e. where is the reformat option) and also the atomic actions required to carry out this task. By constantly dividing the user attention they are likely to remain in *wayfinding* (Svensson 1998) mode. Whilst *wayfinding* in itself is not inherently bad it may take longer and places a greater mental load on the traveller than transportation. This is due to the fact that not only are they having to navigate around within the environment but are also having to deal with minute levels of interaction which may be of minimal if any interest to them. Downs and Stea (Downs and Stea 1973) provide four steps which exist within wayfinding:

1. Orientating oneself in the environment.
2. Choosing the correct route.
3. Monitoring the route.
4. Recognising a destination has been reached.

These areas can broadly be related to GUI's as follows:

1. When the user examines the interface in order to find out where they are in relation to the tools they need in order to complete a task. For example looking round Word in order to find the group of tools related to type face style.
2. In a wizard selecting the correct options to get to the desired end results.
3. The ability to review and check the correct route has been taken. This can partially be achieved in document wizards by providing a stage by stage view of what has been carried out so far.
4. Using a document as an example, providing a view of the finished result. This result should correspond to the users desired intentions.

A system which exhibits support for wayfinding will hopefully also provide strong enough navigational cues for the user to move from *wayfinding* to *transportation* mode as soon as possible.

In order to support the move from a purely atomic interaction several areas pertaining to "being lost in hyperspace" are of potential relevance. Therefore for wayfinding to be adequately supported it is imperative for the user to be aware of their location. A number of hypertext systems already do so by the use of maps, directional links, use of well worn paths and history enriched cues (Svensson 1998) etc. These all provide simple cues which can aid the user in their ability to prevent being "lost in hyperspace" these can all be parallel to areas of GUI design, for example in Wizards buttons such as previous and next are used, which link the various dialogue boxes, in addition some interfaces support the changing of colour of buttons once selected.

Xerox proposed the rooms system, where related applications or facets of information could be grouped into so called "rooms". This concept has been widely adopted in X-Windows type environments through the use of the virtual desktop. The virtual desktop concept falls neatly in to the related field of focus spaces. Focus spaces can apply equally to the physical grouping of related applications or related pieces of information. Focus spaces allow users (or system designers) to layout information in a way which allows them to easily interpret the information, whilst retaining the ability to focus on the areas of interest to them. In essence the idea of a focus space is to allow the user to focus attention on what is important thus reducing problems of information overload and divided attention. The rooms idea also broadly fits in with the idea of on and off screen space in films (Persson 1998) and preservation of context which is discussed to some extent later on in this paper.

Presenting Opportunities, Emergent Behaviour and Extensibility.

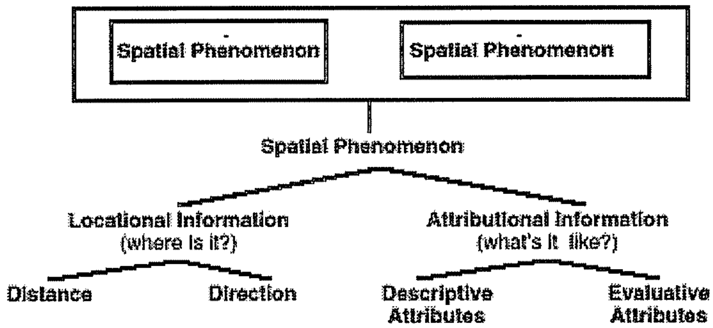


Figure 2: An interface (a spatial phenomenon) consisting of two other spatial phenomenon (e.g. two buttons placed in close proximity).

Utilising space as the underlying principal method of interface design another series of properties also emerge (Erikson 1993). Grouping in essence promotes the idea of spaces for particular tasks and relates to the Lynch concepts. Further to this objects encourage incidental interaction. In essence if the user already has a predefined task, if the interface suggests an appropriate alternative action they may take that option instead. Again using word as an example the **bold,** *italic* and underline operations are all situated together in a group, and the user is presented these three opportunities.

A graphical user interface as has already been highlighted is a spatial environment consisting of the basic layout of components however these considerations take no account of the context in which the user operates. In merely placing objects in an interface we take no account of any emergent properties which may arise either from user or the information. In MUD'S "[the environment is] extensible by the participants", by reducing the ability of the users to adapt the interface we are therefore reducing their ability to interact in a spatial fashion as information and cognitive maps rarely remain static. Extensibility could include the ability to provide functions to allow aspects such as overview maps, annotations (e.g. pop-up context help written by the user) or the ability to redesign the interface. These all parallel with aspects of hypermedia systems which often provide the ability to generate information maps (Baldonado and Winograd 1997) and annotations. Research (Andrews 1997; Baldonado and Winograd 1997) (Card, Robertson et al. 1997) from hypermedia systems would appear to suggest the users ability to customise information is critical to their ability to navigate within the environment. This level of customisation may simply be annotations, or may extend to the ability to re-organise the way in which the knowledge is structured (from the users point of view).

Providing the user with the ability to extend or amend the interface draws into the next area the need for user/system defined focus spaces. The issue of need to support focused and divided attention has already been mentioned to some extent. Novice users frequently encounter problems with information overload both in graphical user interfaces (Nielsen 1993) and in hypermedia systems. However the use of focus spaces can aid in reducing this problem. Focus spaces are somewhat related to the issue of grouping which has already been mentioned.

Transparent Interfaces, Automaticity and Mental Models

The discussion thus far has primarily existed on the level of abstracting the interface from the information environment. However as has been indicated earlier on and in other studies, if the user has to grapple not only with the interface but also the information, there is a risk of introducing divided attention. In order to reduce the problem of divided attention research has been carried out into transparent interfaces (Harrison, Ishii et al. 1995) (Zhai, Buxton et al. 1995) such as menus (Harrison and Vicente 1996) and toolglasses (Bier, Stone et al. 1993).This it was hoped would reduce the amount of divided attention suffered by users, as they no longer have to look from information to application control tools (e.g. in Word). Also by introducing the transparent toolglass system proposed, a number of other issues could also be addressed. Firstly the toolglass system could include some form of contextual awareness. The behaviour of the tools, and layout would change in accordance with where the toolglass was positioned on screen. This allows interfaces to become contextually aware. In film based environments although the viewer only sees the given camera angle they are aware of the context in which that particular scene or shot is placed (Persson 1998) .The screen display on a computer system sits in part in a wider working context. This may be the relationship between different pieces of information on screen or the real world environment. The toolglass can be thought of as the screen which sits within the wider environment, where the tools sit over a specific information artefact. The tools then adapt the information or adapt to the information. This adaption takes place within the wider environmental context. It should be noted that transparency can be thought of as existing on two levels, physical (as in the case of the toolglass) and metaphorical. At the metaphorical level transparency can be said to exist in the MS-Word 6 interface, where by using a right mouse click a contextually aware pop-up menu appears. Further to both the contextual awareness of the tools and reducing divided attention the issue of screen clutter can also be partially resolved.

Following on from the issue of interface transparency, research (Altman, Larkin et al. 1995) has indicated the difficulty experienced by users when using scroll bars to navigate within large documents. Clearly this is unsuitable as the entire point of the user interface is to permit them to work with the information. This can perhaps equally apply to the problem of them retaining mental models of how to complete a given task. Therefore the interface structure should permit the ability of the user to retain the information they work with as well as how to complete the task. In essence therefore the emphasis is on minimising cognitive load by allowing the activity to become (Norman 1986) automatic. Automatic activity (automoticity) arises when a user is familiar with a task and can abstract the level of interaction to the goal of completing that task, where as with less familiar tasks where the abstraction is at a very granular level. Norman's model (Norman 1986) can in part be paralleled with the idea of travellers who initially use route knowledge then eventually through time start using survey knowledge. As has already been indicated we are in essence trying to get users to the use of survey knowledge as quickly as possible, thus lessening the time taken for the task to become automatic. However as can be seen in Sjolinder (Sjölinder 1998) users are a diverse group of people and a number of individual differences arise such as age and sex. Therefore other non-spatial forms of navigational support should where possible also be included.

The idea of automaticity fits in with the concept of allowing travellers to *transport* themselves to a given location and also with the "black box in a glass box" (Höök 1996) view of hypermedia. The "black box in a glass box" approach is aimed at reducing information overload to the user in order to prevent users from getting "lost in hyperspace". This concept can perhaps also broadly be applied to GUI's where "we hide the complex behaviour of the adaptive system in a black box and show a fairly simple model to the user in a glass box". Adopting this idea in GUI's we could perhaps see the interface as being adaptive to some extent. The glass box approach can be applied to both the metaphorical and physical aspects of interface transparency. Firstly we provide enough information as to allow the user to gauge what the tool does, secondly however only a limited amount of information is presented. In doing the latter we seek to prevent overloading the user with information and thus resulting in them getting lost in the interface.

In order for users to commit to memory the mental models of information and tasks, it is important that the number of steps required to complete a task are kept to a minimum. Therefore this raises the issues again of direct manipulation and see and point both of which encourage large numbers of atomic interactions. If we are to assume that the user has no interest in certain tasks (Gentner and Neilsen 1996), then clearly some form of intelligent interface agents may be of value. These agents could take over the running of routine tasks or highly complex but non critical tasks, however there would be a need to retain a high degree of transparency from the users point of view.

Consistency

Consistency in a user interface is widely regarded as an aspect which aids the user ability to learn a piece of software or environment. An example of how this is achieved can be found in Netscape where a single interface style allows interaction with various different types of information. Consistency also applies to real world navigation systems such as road and tourist information signs. An excellent example of easy to understand signs which are laid out according to what is appropriate to usage and work patterns can be found at Amsterdam Schipol airport. Here the language used in the signs is very clear allowing speakers of different languages to understand the meaning. In addition they are positioned at various locations within the buildings most likely to be of benefit to the travellers and with minimal interference to the rest of the environment. Although the signs may be physically located in different locations in the environment, they often adhere to certain layout and style guidelines. The approach adopted by town planners to consistency is one which could be used within computer environments. For example consistency exists in the style of signs, however their positioning is related to the effect it will have on the cosmetic and work patterns of the environment in which they are situated. Consistency is a contested subject while advocates it others such as Gentner (Gentner and Neilsen 1996) would argue that a dynamic environment which is tied down with static guidelines may be unsuitable in some respects for the tasks it were intended. Therefore whilst traditional GUI guideline suggest consistency as being a key component, in order to promote effective navigation within the environment perhaps it should be relaxed in order to allow the use of more effective navigational cues.

Supporting Navigation in GUI's

As already mentioned there are four broad categories which can apply to the list of navigational assessment criteria which follow later on in this paper. However in many cases the assessment criteria applied to more than one of the four. The latter two areas (contextual awareness and support for user behaviour) are largely supersets of the support for physical navigation guidelines. The criteria listed are based on the background information review. The following lists consists of a number of points which are found in traditional HCI methods. In addition however here the focus is on adapting them where possible to support navigation. For example supporting task based interaction can be in part paralleled with the idea of transportation in real world environments. For example a traveller with experience of a specific geographical area has a general view of where they want to go. A traveller with such knowledge can do so without having to concern themselves with the very minute physical details such as placing one foot in front of the other and can focus instead on going in the desired direction.

Support for Physical Navigation

- Strong set of navigational Cues.
    - use of landmarks, paths and districts in GUI.
    - use of colour and suitable icons etc.
- Adhere to certain level of consistency. However consistency should not be at the expense of loss of strong navigational cues.
    - e.g. users of Netscape can in the main interact with a variety if different data types and structures by using a similar or identical interface.
- Transparent interface both to the user and also in terms of the required task.
- Use a rich language for communication both to and from the computer.
    - an example of a basic widget being enhanced is the ability of newer applications scroll bars to be moved up and down by using right or left mouse button. Rather than using the arrows positioned at the bottom of the scroll bar.
- Support Incidental activity.
    - group related icons/menu options near each other.
- Allow user to customise the interface in various ways
    - support customisation of interface layout.
    - annotations in the interface (e.g. user defined pop-up context based help).
    - support for user defined wizards.

Support for Learning Survey Knowledge

- Increase the ability for the task to become automatic thus speeding up the time taken for the transfer from route to survey knowledge.
- Support for task based interaction.
    - make it clear when a task has been completed (a destination has been reached).
- Provision of maps and other navigational support structures within the interface.

Support for Contextual Awareness

- Provide adaptive interfaces which change based on the context of the user or information.
    - e.g. pop-up menus in Word or Internet Explorer.
- Support user in constructing interfaces which are based on their working context.
    - e.g. Xerox Rooms or similar for organising applications into groups.
- Support users ability to retain information.
- scroll bars used in large documents do not support this as users frequently have problems retaining any awareness of the structure of the document.

Support for User Behaviour

- Support for focus spaces (focused and divided attention).
- Reduce need for performance of routine and complex tasks.
- Support customisation of interface and information.
    - e.g. short cut keys.
- Support for task based interaction.
- Support for emergent interest in the aspects of the interface or information.

Conclusion

As is evident from the previous sections in this paper the focus has primarily been on supporting spatial awareness within GUI systems. However this is not always the most appropriate way of providing navigation within environments as many potential users may have poor spatial ability (Sjölinder 1998) . Therefore further work is required in order to examine non-spatial methods of supporting navigation within computer environments. Further to this it is evident that adaptive systems may provide methods of enhancing the users ability to navigate spatially however it is clear from some research (Höök and Svensson 1998) that users frequently have problems with adaptive systems. In summary therefore supporting navigation where the aim is to get from route to survey knowledge as quickly as possible whilst also supporting the move from wayfinding to transportation may require certain trade-offs.

References

Altman, E., J. H. Larkin, et al. (1995). Display Navigation By An Expert Programmer: A Preliminary Model of Memory. ACM Proceedings of Computer and Human Interaction 95.

Andrews, K. (1997). Navigational Facilities for Large Structured Web Spaces. http://www.iicm.edu/heith/chi97.

Baldonado, M. Q. W. and T. Winograd (1997). SenseMaker: An Information-Exploration Interface Supporting the Contextual Evolution of a Users Interests. Computer and Human Interaction 97, ACM.

Bier, E. A., M. C. Stone, et al. (1993). Toolglass and Magic Lenses: The See-Through Interface. SIGGRAPH, Computer Graphics Annual Conference Series, Anaheim, Germany, ACM.

Card, S. K., G. G. Robertson, et al. (1997). The WebBook and the Web Forager: An Information Workspace for the World-Wide Web. Computer and Human Interaction 97, ACM.

Dieberger, A. (1996). Browsing the WWW by Interacting with a textual virtual environment -- A Framework for experimenting with navigational metaphors. Hypertext '96, Washington, ACM.

Dieberger, A. and J. G. Tromp (1993). The Information City - A Metpahor for Navigating in Hypertexts.

Downs, R. and D. Stea (1973). Cognitive Representations. Image and Environment. R. Downs and D. Stea. Chicago, Adline.

Erikson, T. (1993). "From Interface to Interplace: The Spatial Environment as a Medium for Interaction." Proceedings of Conference on Spatial Information Theory, 1993.

Gentner, D. and J. Neilsen (1996). "The Anit-Mac Interface." Communications of the ACM **http://www.acm.org/cacm/AUG96/antimac.htm**.

Harrison, B. L., H. Ishii, et al. (1995). Transparent Layered User Interfaces: An Evaluation of a Display Design to Enhance Focussed and Divided Attention. Computer Human Interaction, Denver, ACM.

Harrison, B. L. and K. J. Vicente (1996). An Experiemental Evaluation of Transparent Menu Usage. Computer and Human Interaction '96, Vancouver, ACM.

Höök, K. (1996). A Glass Box Approach to Adaptive Hypermedia. Stockholm, University of Stockholm and Swedish Institute of Computer Science.

Höök, K. and M. Svensson (1998). Evaluating Adaptive Navigation Support (Position Paper). *This Volume*, SICS.

Lynch, K. (1982). Image of the City, MIT Press.

Nielsen, J. (1993). "Noncommand User Interfaces." Communications of the ACM **36**(4): 83-99.

Norman, D. A. (1986). User Centered Systems Design: New Perspectives on Human-Computer Interaction, Lawrence Erlbaum Associates.

Persson, P. (1998). A Comparative Study of Digital and Cinematic Space with Special Focus on Navigational Issues. *This Volume*, Swedish Institute of Computer Science.

Shum, S. (1990). "Real and Virtual Spaces: Mapping From Spatial Cognition to Hypertext." Hypermedia **2**(2): 133-158.

Sjölinder, M. (1998). Individual differences in spatial cognition and hypermedia navigation. *This Volume*, Swedish Institute of Computer Science.

Streitz, N., J. Haake, et al. (1992). SEPIA: A Cooperative Hypermedia Authoring Environment. Hypertext, Milan, Italy, ACM.

Svensson, M. (1998). Social Navigation. *This volume.*, Swedish Institute of Computer Science.

Zhai, S., W. Buxton, et al. (1995). The Partial Occlusion Effect: Utilising Semi Transparancy in 3D Human Computer Interaction. http://vered.rose.utoronto.ca/people/shumin_dit/silk.htm.